GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 7, 2004, 04:47:28 ; Search time 4615 Seconds
(without alignments)
2090.379 Million cell updates/sec

Title: US-10-071-174-2
Perfect score: 1084
Sequence: 1 MVDQLRERTTMADPLRERTE.....QAFLSCLLTTAFIYLWTRLL 204

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10071174/runat_05112004_131903_23984/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10071174_@CGN_1_1_7406_@runat_05112004_131903_23984 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1084	100.0	615	9	AF285092	AF285092 Homo sapi
	2	1084	100.0	887	9	AF326964	AF326964 Homo sapi
	3	1045	96.4	1168	6	BD233466	BD233466 Human pro
	4	1036.5	95.6	726	6	CQ752105	CQ752105 Sequence
	5	1035	95.5	582	6	BD233456	BD233456 Human pro
	6	1035	95.5	585	9	HSA458330	AJ458330 Homo sapi
c	7	855	78.9	93287	9	AC023906	AC023906 Homo sapi
	8	494	45.6	214669	2	AC018903	AC018903 Homo sapi
	9	487	44.9	1074	10	AY029163	AY029163 Rattus no
	10	474.5	43.8	1209	10	AF102501	AF102501 Mus muscu
	11	474.5	43.8	1225	10	AF067660	AF067660 Mus muscu
	12	467.5	43.1	1257	10	BC052690	BC052690 Mus muscu
	13	342	31.5	240461	2	AC111669	AC111669 Rattus no
	14	333.5	30.8	169914	2	AC133947	AC133947 Mus muscu
	15	333.5	30.8	202851	10	AC115880	AC115880 Mus muscu
c	16	218.5	20.2	214669	2	AC018903	AC018903 Homo sapi
	17	193.5	17.9	1475	5	CCNR13	X84418 C.coturnix
	18	188.5	17.4	534	5	AF375661	AF375661 Gallus ga
	19	179	16.5	531	5	AF441285	AF441285 Danio rer
c	20	178.5	16.5	151102	14	AF282130	AF282130 Meleagrid
	21	178.5	16.5	159160	14	AF291866	AF291866 Meleagrid
c	22	178.5	16.5	159160	14	AF291866	AF291866 Meleagrid
c	23	144	13.3	170193	5	AL935189	AL935189 Zebrafish
	24	143.5	13.2	636	5	AF120210	AF120210 Gallus ga
	25	142.5	13.1	712	5	CR390222	CR390222 Gallus ga
	26	135.5	12.5	9471	5	AY186728	AY186728 Gallus ga
	27	132	12.2	243686	2	AC096430	AC096430 Rattus no
	28	131	12.1	694	4	AB080230	AB080230 Canis fam
	29	129	11.9	658	6	AX525910	AX525910 Sequence
	30	127	11.7	558	9	HSA417988	AJ417988 Homo sapi
	31	127	11.7	579	6	E58778	E58778 Screening m
	32	127	11.7	579	6	AR380842	AR380842 Sequence
	33	127	11.7	579	6	AX057142	AX057142 Sequence
	34	127	11.7	579	6	AX481419	AX481419 Sequence
	35	127	11.7	579	9	HUMBAXA	L22473 Human Bax a
	36	127	11.7	779	6	BD215262	BD215262 Novel hum
	37	127	11.7	791	6	BD203742	BD203742 Human nuc
	38	127	11.7	791	6	AX014897	AX014897 Sequence
	39	127	11.7	871	9	BC014175	BC014175 Homo sapi
	40	127	11.7	996	6	AR429416	AR429416 Sequence
	41	126	11.6	646	4	AB080724	AB080724 Felis cat
	42	126	11.6	677	9	AF007826	AF007826 Homo sapi
	43	125	11.5	525	5	AF120211	AF120211 Gallus ga
	44	125	11.5	624	6	AR027309	AR027309 Sequence
	45	125	11.5	624	6	I76309	I76309 Sequence 1

ALIGNMENTS

RESULT 1
AF285092
LOCUS AF285092 615 bp mRNA linear PRI 08-NOV-2001
DEFINITION Homo sapiens Bcl-2-like protein 10 mRNA, complete cds.
ACCESSION AF285092
VERSION AF285092.1 GI:9837265
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 615)
AUTHORS Zhang,H., Holzgreve,W. and De Geyter,C.
TITLE Bcl2-L-10, a novel anti-apoptotic member of the Bcl-2 family,
blocks apoptosis in the mitochondria death pathway but not in the
death receptor pathway
JOURNAL Hum. Mol. Genet. 10 (21), 2329-2339 (2001)
MEDLINE 21548034
PUBMED 11689480
REFERENCE 2 (bases 1 to 615)
AUTHORS Zhang,H.H.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-2000) University Women's Hospital, Schanzenstra
46, Basel 4057, Switzerland
FEATURES Location/Qualifiers
source 1. .615
/organism="Homo sapiens"
/mol_type="mRNA"

```
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21"
                     /tissue_type="ovary"
     CDS             1. .615
                     /codon_start=1
                     /product="Bcl-2-like protein 10"
                     /protein_id="AAG00503.1"
                     /db_xref="GI:9837266"
                     /translation="MVDQLRERTTMADPLRERTELLLADYLGYCAREPGTPEPAPSTP
                     EAAVLRSAAARLRQIHRSFFSAYLGYPGNRFELVALMADSVLSDSPGPTWGRVVTLVT
                     FAGTLLERGPLVTARWKKWGFQPRLKEQEGDVARDCQRLVALLSSRLMGQHRAWLQAQ
                     GGWDGFCHFFRTPFPLAFWRKQLVQAFLSCLLTTAFIYLWTRLL"
ORIGIN

Alignment Scores:
Pred. No.:                4.47e-81        Length:          615
Score:                    1084.00         Matches:         204
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-10-071-174-2 (1-204) x AF285092 (1-615)

Qy          1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 60

Qy         21 LeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArgGluProGlyThrProGluProAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 120

Qy         41 ProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHis 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 180

Qy         61 ArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 240

Qy         81 MetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 300

Qy        101 ValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 360

Qy        121 TrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 420

Qy        141 LeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGln 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 480

Qy        161 GlyGlyTrpAspGlyPheCysHisPhePheArgThrProPheProLeuAlaPheTrpArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGA 540

Qy        181 LysGlnLeuValGlnAlaPheLeuSerCysLeuLeuThrThrAlaPheIleTyrLeuTrp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGG 600

Qy        201 ThrArgLeuLeu 204
              ||||||||||||
Db        601 ACACGATTATTA 612

RESULT 2
AF326964
LOCUS       AF326964                887 bp    mRNA    linear   PRI 01-MAY-2001
DEFINITION  Homo sapiens BCLB (BCLB) mRNA, complete cds.
ACCESSION   AF326964
VERSION     AF326964.1  GI:13898393
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 887)
  AUTHORS   Ke,N., Godzik,A. and Reed,J.C.
  TITLE     Bcl-B, a novel Bcl-2 family member that differentially binds and
            regulates Bax and Bak
  JOURNAL   J. Biol. Chem. 276 (16), 12481-12484 (2001)
  MEDLINE   21201065
   PUBMED   11278245
REFERENCE   2  (bases 1 to 887)
  AUTHORS   Ke,N., Godzik,A. and Reed,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-DEC-2000) The Burnham Institute, 10901 N. Torrey
            Pines Rd., La Jolla, CA 92037, USA
FEATURES             Location/Qualifiers
     source          1. .887
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21"
     gene            1. .887
                     /gene="BCLB"
     CDS             50. .664
                     /gene="BCLB"
                     /note="BCL2 family protein; Boo-like"
                     /codon_start=1
                     /product="BCLB"
                     /protein_id="AAK48715.1"
                     /db_xref="GI:13898394"
                     /translation="MVDQLRERTTMADPLRERTELLLADYLGYCAREPGTPEPAPSTP
                     EAAVLRSAAARLRQIHRSFFSAYLGYPGNRFELVALMADSVLSDSPGPTWGRVVTLVT
                     FAGTLLERGPLVTARWKKWGFQPRLKEQEGDVARDCQRLVALLSSRLMGQHRAWLQAQ
                     GGWDGFCHFFRTPFPLAFWRKQLVQAFLSCLLTTAFIYLWTRLL"
ORIGIN

Alignment Scores:
Pred. No.:                6.57e-81        Length:          887
Score:                    1084.00         Matches:         204
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              100.00%         Indels:          0
DB:                       9               Gaps:            0

US-10-071-174-2 (1-204) x AF326964 (1-887)

Qy          1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 109

Qy         21 LeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArgGluProGlyThrProGluProAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 169

Qy         41 ProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHis 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 229

Qy         61 ArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 289

Qy         81 MetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 349

Qy        101 ValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        350 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 409

Qy        121 TrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 469

Qy        141 LeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGln 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 529

Qy        161 GlyGlyTrpAspGlyPheCysHisPhePheArgThrProPheProLeuAlaPheTrpArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 GGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGA 589

Qy        181 LysGlnLeuValGlnAlaPheLeuSerCysLeuLeuThrThrAlaPheIleTyrLeuTrp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        590 AAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGG 649

Qy        201 ThrArgLeuLeu 204
              ||||||||||||
Db        650 ACACGATTATTA 661
```

```
RESULT 3
BD233466
LOCUS       BD233466                1168 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION   BD233466
VERSION     BD233466.1  GI:33043236
KEYWORDS    JP 2002519016-A/12.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1168)
  AUTHORS   Kato,S. and Kimura,T.
  TITLE     Human protein having hydrophobic domain and DNA encoding the same
  JOURNAL   Patent: JP 2002519016-A 12 02-JUL-2002;
            SAGAMI CHEMICAL RESEARCH CENTER,PROTEGENE INC
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002519016-A/12
            PD   02-JUL-2002
            PF   18-JUN-1999 JP 2000557267
            PI   SEISHI KATO,TOMOKO KIMURA
            PC
            C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
             00
            CC   Human protein having hydrophobic domain and DNA encoding the
            CC           same
            FH   Key             Location/Qualifiers
            FT   source          1. .1168
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1168
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.51e-77        Length:         1168
Score:                  1045.00         Matches:        196
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            96.40%          Indels:         0
DB:                     6               Gaps:           0

US-10-071-174-2 (1-204) x BD233466 (1-1168)

Qy          9 ThrThrMetAlaAspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGly 28
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGG 60

Qy         29 TyrCysAlaArgGluProGlyThrProGluProAlaProSerThrProGluAlaAlaVal 48
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTG 120

Qy         49 LeuArgSerAlaAlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeu 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTC 180

Qy         69 GlyTyrProGlyAsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAsp 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGAC 240

Qy         89 SerProGlyProThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeu 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTG 300

Qy        109 GluArgGlyProLeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLys 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAG 360

Qy        129 GluGlnGluGlyAspValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArg 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGG 420

Qy        149 LeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHis 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCAC 480

Qy        169 PhePheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeu 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAAACAGCTGGTCCAGGCTTTTCTG 540

Qy        189 SerCysLeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeuLeu 204
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 588
```

```
RESULT 4
CQ752105
LOCUS       CQ752105                726 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 38039 from Patent WO02068579.
ACCESSION   CQ752105
VERSION     CQ752105.1  GI:42387450
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 38039 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              4.66e-77        Length:         726
Score:                  1036.50         Matches:        203
Percent Similarity:     89.04%          Conservative:   0
Best Local Similarity:  89.04%          Mismatches:     1
Query Match:            95.62%          Indels:         24
DB:                     6               Gaps:           1

US-10-071-174-2 (1-204) x CQ752105 (1-726)

Qy          1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 60
```

```
Qy          21 LeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArgGluProGlyThrProGluProAla 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CGGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 120

Qy          41 ProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHis 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 180

Qy          61 ArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeu 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 240

Qy          81 MetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLeu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 300

Qy         101 ValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLys 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 360

Qy         121 TrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArg 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 420

Qy         141 LeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGln 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 480

Qy         161 GlyGlyTrp------------------------------------------------------ 163
               |||||||||
Db         481 GGCGGCTGGGTGAGCACGCGGCGGACACCGGGACACGGGGCGGGACGGGCAGCCGGGAAG 540

Qy         164 -------------------AspGlyPheCysHisPhePheArgThrProPheProLeuAl 177
                                  ||||||||||||||||||||||||||||||||||||||||
Db         541 CGCCCACGAGGCTGGCACGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGC 600

Qy         177 aPheTrpArgLysGlnLeuValGlnAlaPheLeuSerCysLeuLeuThrThrAlaPheIl 197
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 TTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCAT 660

Qy         197 eTyrLeuTrpThrArgLeuLeu 204
               |||||||||||||||||||||
Db         661 TTATCTCTGGACACGATTATTA 682
```

RESULT 5
BD233456

```
LOCUS       BD233456                582 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION   BD233456
VERSION     BD233456.1  GI:33043226
KEYWORDS    JP 2002519016-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Kato,S. and Kimura,T.
  TITLE     Human protein having hydrophobic domain and DNA encoding the same
  JOURNAL   Patent: JP 2002519016-A 2 02-JUL-2002;
            SAGAMI CHEMICAL RESEARCH CENTER,PROTEGENE INC
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002519016-A/2
            PD   02-JUL-2002
            PF   18-JUN-1999 JP 2000557267
            PI   SEISHI KATO,TOMOKO KIMURA
            PC
            C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
             00
            CC   Human protein having hydrophobic domain and DNA encoding the
            CC             same
            FH   Key             Location/Qualifiers
            FT   source          1. .582
            FT                   /organism='Homo sapiens (human)'.
FEATURES    Location/Qualifiers
     source 1. .582
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:                4.92e-77      Length:          582
Score:                    1035.00       Matches:         194
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              95.48%        Indels:          0
DB:                       6             Gaps:            0

US-10-071-174-2 (1-204) x BD233456 (1-582)

Qy          11 MetAlaAspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCys 30
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 60

Qy          31 AlaArgGluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArg 50
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 120

Qy          51 SerAlaAlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyr 70
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 180

Qy          71 ProGlyAsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerPro 90
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 240

Qy          91 GlyProThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArg 110
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 300

Qy         111 GlyProLeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGln 130
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 360

Qy         131 GluGlyAspValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeuMet 150
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 420

Qy         151 GlyGlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHisPhePhe 170
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 480

Qy         171 ArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeuSerCys 190
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 540

Qy         191 LeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeuLeu 204
               ||||||||||||||||||||||||||||||||||||||||||
Db         541 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 582
```

RESULT 6
HSA458330

```
LOCUS       HSA458330               585 bp    DNA     linear   PRI 25-APR-2002
DEFINITION  Homo sapiens NRH gene for anti-apoptotic protein.
ACCESSION   AJ458330
VERSION     AJ458330.1  GI:20338765
KEYWORDS    anti-apoptotic protein; NRH gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Aouacheria,A., Arnaud,E., Venet,S., Lalle,P., Gouy,M., Rigal,D. and
            Gillet,G.
  TITLE     Nrh, a human homologue of Nr-13 associates with Bcl-Xs and is an
```

```
          inhibitor of apoptosis
  JOURNAL   Oncogene 20 (41), 5846-5855 (2001)
  MEDLINE   21477277
   PUBMED   11593390
REFERENCE   2  (bases 1 to 585)
  AUTHORS   Gillet,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-2002) Gillet G., Ibcp, CNRS UMR 5086, 7 PASSAGE
            DU VERCORS, 69367, FRANCE
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
     gene            1. .585
                     /gene="NRH"
     CDS             1. .585
                     /gene="NRH"
                     /function="apoptosis inhibitor"
                     /codon_start=1
                     /product="anti-apoptotic protein"
                     /protein_id="CAD30221.1"
                     /db_xref="GI:20338766"
                     /db_xref="GOA:Q9HD36"
                     /db_xref="Swiss-Prot:Q9HD36"
                     /translation="MADPLRERTELLLADYLGYCAREPGTPEPAPSTPEAAVLRSAAA
                     RLRQIHRSFFSAYLGYPGNRFELVALMADSVLSDSPGPTWGRVVTLVTFAGTLLERGP
                     LVTARWKKWGFQPRLKEQEGDVARDCQRLVALLSSRLMGQHRAWLQAQGGWDGFCHFF
                     RTPFPLAFWRKQLVQAFLSCLLTTAFIYLWTRLL"
ORIGIN

Alignment Scores:
Pred. No.:                4.95e-77      Length:          585
Score:                    1035.00       Matches:         194
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              95.48%        Indels:          0
DB:                       9             Gaps:            0

US-10-071-174-2 (1-204) x HSA458330 (1-585)

Qy          11 MetAlaAspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCys 30
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 60

Qy          31 AlaArgGluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArg 50
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 120

Qy          51 SerAlaAlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyr 70
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 180

Qy          71 ProGlyAsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerPro 90
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 240

Qy          91 GlyProThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArg 110
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 300

Qy         111 GlyProLeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGln 130
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 360

Qy         131 GluGlyAspValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeuMet 150
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 420

Qy         151 GlyGlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHisPhePhe 170
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 480

Qy         171 ArgThrProProProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeuSerCys 190
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 540

Qy         191 LeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeuLeu 204
               |||||||||||||||||||||||||||||||||||||||||||
Db         541 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 582

RESULT 7
AC023906/c
LOCUS       AC023906               93287 bp    DNA     linear   PRI 06-JUL-2001
DEFINITION  Homo sapiens chromosome 15 clone CTD-2184D3 map 15q21.2, complete
            sequence.
ACCESSION   AC023906
VERSION     AC023906.7  GI:14595770
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2000) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
REFERENCE   4  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
COMMENT     On Jul 4, 2001 this sequence version replaced gi:12248292.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399

            Note:  Data from overlapping BACs AC010674 [drafting center:
            UWMSC], AC090970 [drafting center: UWMSC], and AC016824 [drafting
            center: GTC] was added for finishing
            ------------.
FEATURES             Location/Qualifiers
     source          1. .93287
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
```

```
                     /map="15q21.2"
                     /clone="CTD-2184D3"
                     /clone_lib="Cal Tech Human BAC library D"
                     /note="Data from overlapping clones CTD-2650P22 AC090970,
                     RP11-47K1 AC016824, and RP11-430B1 was added and the
                     consensus was determined from CTD-2184D3 to the extent
                     possible"
     misc_feature    1. .8893
                     /note="overlap with CTD-2650P22 AC090970"
     misc_feature    1. .8893
                     /note="overlap with RP11-47K1 AC016824"
     unsure          6078. .6140
                     /note="low quality data"
     unsure          10529. .10532
                     /note="low quality data"
     unsure          14722. .14808
                     /note="low quality data"
     unsure          37627
                     /note="low quality data"
     unsure          71552. .71593
                     /note="low quality data"
     misc_feature    82012. .93287
                     /note="overlap with RP11-430B1 AC010674"
ORIGIN

Alignment Scores:
Pred. No.:              8.92e-60        Length:         93287
Score:                  855.00          Matches:        163
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            78.87%          Indels:         0
DB:                     9               Gaps:           0

US-10-071-174-2 (1-204) x AC023906 (1-93287)

Qy          1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64942 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 64883

Qy         21 LeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArgGluProGlyThrProGluProAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64882 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 64823

Qy         41 ProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHis 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64822 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 64763

Qy         61 ArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64762 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 64703

Qy         81 MetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64702 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 64643

Qy        101 ValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64642 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 64583

Qy        121 TrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64582 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 64523

Qy        141 LeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGln 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64522 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 64463

Qy        161 GlyGlyTrp 163
              |||||||||
Db      64462 GGCGGCTGG 64454

RESULT 8
AC018903
```

```
LOCUS       AC018903               214669 bp    DNA     linear   HTG 04-JUN-2000
DEFINITION  Homo sapiens chromosome 15 clone RP11-337B11 map 15q21, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC018903
VERSION     AC018903.2  GI:8247797
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 214669)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., Kaur,A.,
            Madan,A., Nesbitt,R., Shaffer,T. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 214669)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-1999) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
COMMENT     On Jun 4, 2000 this sequence version replaced gi:6630517.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399
            ------------.
            * NOTE: This record contains 192 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      806: contig of 806 bp in length
            *      807      906: gap of unknown length
            *      907     2068: contig of 1162 bp in length
            *     2069     2168: gap of unknown length
            *     2169     3020: contig of 852 bp in length
            *     3021     3120: gap of unknown length
            *     3121     4284: contig of 1164 bp in length
            *     4285     4384: gap of unknown length
            *     4385     5201: contig of 817 bp in length
            *     5202     5301: gap of unknown length
            *     5302     6544: contig of 1243 bp in length
            *     6545     6644: gap of unknown length
            *     6645     7472: contig of 828 bp in length
            *     7473     7572: gap of unknown length
            *     7573     8756: contig of 1184 bp in length
            *     8757     8856: gap of unknown length
            *     8857     9658: contig of 802 bp in length
            *     9659     9758: gap of unknown length
            *     9759    10905: contig of 1147 bp in length
            *    10906    11005: gap of unknown length
            *    11006    11811: contig of 806 bp in length
            *    11812    11911: gap of unknown length
            *    11912    13293: contig of 1382 bp in length
            *    13294    13393: gap of unknown length
            *    13394    14211: contig of 818 bp in length
            *    14212    14311: gap of unknown length
```

```
*    14312    15465: contig of 1154 bp in length
*    15466    15565: gap of unknown length
*    15566    16379: contig of 814 bp in length
*    16380    16479: gap of unknown length
*    16480    17624: contig of 1145 bp in length
*    17625    17724: gap of unknown length
*    17725    18578: contig of 854 bp in length
*    18579    18678: gap of unknown length
*    18679    19774: contig of 1096 bp in length
*    19775    19874: gap of unknown length
*    19875    20700: contig of 826 bp in length
*    20701    20800: gap of unknown length
*    20801    22036: contig of 1236 bp in length
*    22037    22136: gap of unknown length
*    22137    23023: contig of 887 bp in length
*    23024    23123: gap of unknown length
*    23124    24196: contig of 1073 bp in length
*    24197    24296: gap of unknown length
*    24297    25141: contig of 845 bp in length
*    25142    25241: gap of unknown length
*    25242    26385: contig of 1144 bp in length
*    26386    26485: gap of unknown length
*    26486    27325: contig of 840 bp in length
*    27326    27425: gap of unknown length
*    27426    28691: contig of 1266 bp in length
*    28692    28791: gap of unknown length
*    28792    29655: contig of 864 bp in length
*    29656    29755: gap of unknown length
*    29756    30881: contig of 1126 bp in length
*    30882    30981: gap of unknown length
*    30982    31813: contig of 832 bp in length
*    31814    31913: gap of unknown length
*    31914    33056: contig of 1143 bp in length
*    33057    33156: gap of unknown length
*    33157    34007: contig of 851 bp in length
*    34008    34107: gap of unknown length
*    34108    35320: contig of 1213 bp in length
*    35321    35420: gap of unknown length
*    35421    36207: contig of 787 bp in length
*    36208    36307: gap of unknown length
*    36308    37521: contig of 1214 bp in length
*    37522    37621: gap of unknown length
*    37622    38420: contig of 799 bp in length
*    38421    38520: gap of unknown length
*    38521    39726: contig of 1206 bp in length
*    39727    39826: gap of unknown length
*    39827    40635: contig of 809 bp in length
*    40636    40735: gap of unknown length
*    40736    41845: contig of 1110 bp in length
*    41846    41945: gap of unknown length
*    41946    42730: contig of 785 bp in length
*    42731    42830: gap of unknown length
*    42831    43917: contig of 1087 bp in length
*    43918    44017: gap of unknown length
*    44018    44825: contig of 808 bp in length
*    44826    44925: gap of unknown length
*    44926    46114: contig of 1189 bp in length
*    46115    46214: gap of unknown length
*    46215    47031: contig of 817 bp in length
*    47032    47131: gap of unknown length
*    47132    48292: contig of 1161 bp in length
*    48293    48392: gap of unknown length
*    48393    49176: contig of 784 bp in length
*    49177    49276: gap of unknown length
*    49277    50389: contig of 1113 bp in length
*    50390    50489: gap of unknown length
*    50490    51333: contig of 844 bp in length
*    51334    51433: gap of unknown length
*    51434    52765: contig of 1332 bp in length
*    52766    52865: gap of unknown length
*    52866    53710: contig of 845 bp in length
*    53711    53810: gap of unknown length
*    53811    54965: contig of 1155 bp in length
*    54966    55065: gap of unknown length
*    55066    55910: contig of 845 bp in length
*    55911    56010: gap of unknown length
*    56011    57558: contig of 1548 bp in length
*    57559    57658: gap of unknown length
*    57659    58509: contig of 851 bp in length
*    58510    58609: gap of unknown length
*    58610    59745: contig of 1136 bp in length
*    59746    59845: gap of unknown length
*    59846    60655: contig of 810 bp in length
*    60656    60755: gap of unknown length
*    60756    61976: contig of 1221 bp in length
*    61977    62076: gap of unknown length
*    62077    62896: contig of 820 bp in length
*    62897    62996: gap of unknown length
*    62997    64128: contig of 1132 bp in length
*    64129    64228: gap of unknown length
*    64229    65026: contig of 798 bp in length
*    65027    65126: gap of unknown length
*    65127    66269: contig of 1143 bp in length
*    66270    66369: gap of unknown length
*    66370    67220: contig of 851 bp in length
*    67221    67320: gap of unknown length
*    67321    68425: contig of 1105 bp in length
*    68426    68525: gap of unknown length
*    68526    69322: contig of 797 bp in length
*    69323    69422: gap of unknown length
*    69423    70568: contig of 1146 bp in length
*    70569    70668: gap of unknown length
*    70669    71547: contig of 879 bp in length
*    71548    71647: gap of unknown length
*    71648    73629: contig of 1982 bp in length
*    73630    73729: gap of unknown length
*    73730    74587: contig of 858 bp in length
*    74588    74687: gap of unknown length
*    74688    75780: contig of 1093 bp in length
*    75781    75880: gap of unknown length
*    75881    77196: contig of 1316 bp in length
*    77197    77296: gap of unknown length
*    77297    78651: contig of 1355 bp in length
*    78652    78751: gap of unknown length
*    78752    79557: contig of 806 bp in length
*    79558    79657: gap of unknown length
*    79658    80774: contig of 1117 bp in length
*    80775    80874: gap of unknown length
*    80875    81690: contig of 816 bp in length
*    81691    81790: gap of unknown length
*    81791    82928: contig of 1138 bp in length
*    82929    83028: gap of unknown length
*    83029    83902: contig of 874 bp in length
*    83903    84002: gap of unknown length
*    84003    85193: contig of 1191 bp in length
*    85194    85293: gap of unknown length
*    85294    86149: contig of 856 bp in length

Alignment Scores:
Pred. No.:                1.96e-29        Length:          214669
Score:                    494.00          Matches:         110
Percent Similarity:       77.63%          Conservative:    8
Best Local Similarity:    72.37%          Mismatches:      33
Query Match:              45.57%          Indels:          2
DB:                       2               Gaps:            0

US-10-071-174-2 (1-204) x AC018903 (1-214669)

Qy          1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
              |||||||||  ||||||||||||||||||||||  ||||||||||||||||||||
Db      93262 ATGGTTGACCATTTGCGGGAGCGCACCACCATGTGCGACCCGCTGCGGGAGCGCACCGAG 93321

Qy         21 LeuLeuLeuAlaAspTyrLeuGlyTyr-CysAlaArgGluProGlyThrProGluProAl 40
              |||||||||  |||||||||  ||| |||  |||||||||  ||||||||||||
Db      93322 CTGTTGCTGCGCGACTACCTAGAGTACCTGCCCCCGGGAACCTCGCACCCCCGAGCCGAG 93381
```

```
Qy          40 aProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHi 60
               |||  |||||||||||||||  |||  |||||||||  ||||||||||||  ||
Db       93382 GCCATTCACACCCGAGGCCGCCGAGCTTGGCTCCGCAGCCCGCAGGTTACGGCAGAATCA 93441

Qy          60 sArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLe 80
               |||||||||||||      ||||||||||||||||   |||||||||:::||||||||||
Db       93442 CCGGTCCTTTTTTCTCCCCTACCTCGGATACCCCGAAAACCGCTTCGACCTGGTGGCGCT 93501

Qy          80 uMetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLe 100
               ||||||||||||||||||||||||  |||  ||||||  |||:::      |||||
Db       93502 GATGGCGGATTCCGTACTCTCCGACAGACCCGTGCCCACCTAGGGCAAAAAGGGGACGCT 93561

Qy         100 uValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLy 120
               ||||||||||:::||||||||||||||:::   ||||||||||||        |||||
Db       93562 CGTGACCTTCTCAGGGACACTTCTTGAGAAAAGGCCCCTAGTGACCCCCCCCTTGAAAAA 93621

Qy         120 sTrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnAr 140
               |      ||||||||||||:::|||||||||:::         |||   |||||||||
Db       93622 AAATAGCTTTCAACCCCGGGTAAAAGAACAAAAGAGCACCTGCGCCCA-GACTGCCAACA 93680

Qy         140 gLeuValAlaLeuLeuSerSerArgLeuMetGly 151
               |||   |||||||||:::|||   |||:::|||
Db       93681 CCTGGAAGCCTTACTGAACTCAGAACTTATAGGA 93714
```

RESULT 9
AY029163

LOCUS	AY029163 1074 bp mRNA linear ROD 19-JUN-2003
DEFINITION	Rattus norvegicus BCL2L10 (Bcl2l10) mRNA, complete cds.
ACCESSION	AY029163
VERSION	AY029163.1 GI:13641257
KEYWORDS	.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 1074)
AUTHORS	Itoh,T., Itoh,A. and Pleasure,D.
TITLE	Bcl-2-related protein family gene expression during oligodendroglial differentiation
JOURNAL	J. Neurochem. 85 (6), 1500-1512 (2003)
MEDLINE	22672518
PUBMED	12787069
REFERENCE	2 (bases 1 to 1074)
AUTHORS	Itoh,T., Itoh,A. and Pleasure,D.
TITLE	Direct Submission
JOURNAL	Submitted (29-MAR-2001) Neurology Research, The Children's Hospital of Philadelphia, Abramson Research Center, Room 516 I, 3517 Civic Center Boulevard, Philadelphia, PA 19104, USA

FEATURES Location/Qualifiers

```
     source          1. .1074
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     gene            1. .1074
                     /gene="Bcl2l10"
     CDS             12. .569
                     /gene="Bcl2l10"
                     /note="Boo, Diva"
                     /codon_start=1
                     /product="BCL2L10"
                     /protein_id="AAK31792.1"
                     /db_xref="GI:13641258"
                     /translation="MGDPLQDRTRRLLTDYILFCARAPNTPEPLPTSVEAALLRSVTS
                     QIQQEHQDLFNSFRDYQGNRLELVTQMADELLSNDQEFNWGRLVMLLAFVGTLMNQDR
                     TVKRRRDQRNRLLLERDCYLIVSLLYNRLTGRHRSWLEAHGGWDGFCQFFKNPLPPGF
                     WRRLLIRAILSCFFATAIFYIWKCL"
```

ORIGIN

Alignment Scores:

Pred. No.:	2.84e-31	Length:	1074
Score:	487.00	Matches:	89
Percent Similarity:	64.58%	Conservative:	35
Best Local Similarity:	46.35%	Mismatches:	56
Query Match:	44.93%	Indels:	12
DB:	10	Gaps:	2

US-10-071-174-2 (1-204) x AY029163 (1-1074)

```
Qy          11 MetAlaAspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCys 30
               |||  |||||||||:::::|||||      ||||||  ||||||:::  :::|||
Db          12 ATGGGTGACCCGCTGCAGGATCGCACTAGACGGCTGCTGACTGACTACATATTGTTCTGC 71

Qy          31 AlaArgGluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArg 50
               ||||||  |||  ||||||||||||  |||:::::  |||||||||:::||||||
Db          72 GCACGGGCGCCGAACACCCCTGAGCCACTGCCCACGTCTGTTGAGGCGGCCTTGCTGCGC 131

Qy          51 SerAlaAlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyr 70
               |||      ::::::::::::|||   |||:::      |||:::::::::      |||
Db         132 TCTGTGACTAGTCAGATCCAACAGGAGCACCAGGATCTTTTCAACTCCTTCCGCGACTAC 191

Qy          71 ProGlyAsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerPro 90
               |||||||||  |||||||||      ||||||||   :::||||||:::
Db         192 CAGGGCAACCGCCTGGAGCTGGTGACACAGATGGCGGATGAGTTGCTCTCCAATGACCAA 251

Qy          91 GlyProThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArg 110
                  |||||||||:::|||   |||:::   |||   ||||||||||:::   :::
Db         252 GAGTTCAACTGGGGCCGCCTGGTGATGCTCCTGGCCTTCGTGGGGACGCTAATGAACCAA 311

Qy         111 GlyProLeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGln 130
                  |||      |||                                 ::::::|||
Db         312 GACAGGACTGTTAAGCGGAGG----------------------------AGGGATCAA 341

Qy         131 GluGlyAsp------ValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArg 148
                        :::   |||||||||      :::|||:::||||||   :::|||
Db         342 AGAAACCGTCTCCTACTGGAGCGAGACTGCTATCTCATAGTGAGCTTGCTGTACAATCGA 401

Qy         149 LeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHis 168
               |||  |||:::||||||:::||||||:::|||  ||||||||||||||||||||
Db         402 CTCACAGGACGGCATCGCTCCTGGCTGGAGGCTCACGGTGGCTGGGATGGCTTTTGCCAA 461

Qy         169 PhePheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeu 188
               ||||||:::  |||  |||      |||||||||:::  |||:::::::|||  |||
Db         462 TTCTTCAAGAACCCCTTACCACCCGGCTTCTGGAGAAGATTGCTGATCCGGGCTATTCTG 521

Qy         189 SerCysLeuLeuThrThrAlaPheIleTyrLeuTrp 200
               ||||||      ||||||      |||:::|||
Db         522 TCCTGTTTCTTTGCAACGGCCATCTTTTATATCTGG 557
```

RESULT 10
AF102501

LOCUS	AF102501 1209 bp mRNA linear ROD 21-JAN-1999
DEFINITION	Mus musculus Bcl-2 homolog (Boo) mRNA, complete cds.
ACCESSION	AF102501
VERSION	AF102501.1 GI:4165137
KEYWORDS	.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 1209)
AUTHORS	Song,Q., Kuang,Y., Dixit,V.M. and Vincenz,C.
TITLE	Boo, a novel negative regulator of cell death, interacts with Apaf-1
JOURNAL	EMBO J. 18 (1), 167-178 (1999)
MEDLINE	99094902
PUBMED	9878060
REFERENCE	2 (bases 1 to 1209)
AUTHORS	Song,Q.Z., Kuang,Y.P., Dixit,V.M. and Vincenz,C.
TITLE	Direct Submission
JOURNAL	Submitted (28-OCT-1998) Pathology, University of Michigan, 1301 Catherine Road, Ann Arbor, MI 48109, USA

FEATURES Location/Qualifiers

```
     source          1. .1209
                     /organism="Mus musculus"
```

```
                     /mol_type="mRNA"
                     /strain="B6D2 F1/J"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /map="D9"
                     /tissue_type="ovary"
     gene            1. .1209
                     /gene="Boo"
     CDS             117. .692
                     /gene="Boo"
                     /codon_start=1
                     /product="Bcl-2 homolog"
                     /protein_id="AAD08703.1"
                     /db_xref="GI:4165138"
                     /translation="MADSQDPLHERTRRLLSDYIFFCAREPDTPEPPPTSVEAALLRS
                     VTRQIQQEHQEFFSSFCESRGNRLELVKQMADKLLSKDQDFSWSQLVMLLAFAGTLMN
                     QGPYMAVKQKRDLGNRVIVTRDCCLIVNFLYNLLMGRRHRARLEALGGWDGFCRFFKN
                     PLPLGFWRRLLIQAFLSGFFATAIFFIWKRL"
ORIGIN

Alignment Scores:
Pred. No.:                3.51e-30        Length:          1209
Score:                    474.50          Matches:         94
Percent Similarity:       64.62%          Conservative:    32
Best Local Similarity:    48.21%          Mismatches:      56
Query Match:              43.77%          Indels:          13
DB:                       10              Gaps:            3

US-10-071-174-2 (1-204) x AF102501 (1-1209)

Qy         13 AspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArg 32
              |||||||||   |||||||||       ||||||:::||||||:::   :::|||||||||
Db        132 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 191

Qy         33 GluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArgSerAla 52
              ||||||   ||||||||||||   |||::::::   |||||||||:::||||||||||
Db        192 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 251

Qy         53 AlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyrProGly 72
                  :::::::::|||   |||:::   |||||||||::::::             |||
Db        252 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 311

Qy         73 AsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerProGlyPro 92
              ||||||   |||||||||      ||||||||   :::||||||
Db        312 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 371

Qy         93 ThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArgGlyPro 112
              :::|||   :::::|||   |||:::   |||||||||||||||:::   :::|||||||
Db        372 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 431

Qy        113 LeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGly 132
                 :::      :::   |||:::                                |||
Db        432 TACATGGCTGTCAAGCAGAAGAGG-------------------------GATCTGGGG 464

Qy        133 Asp---------ValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeu 149
              :::         |||   |||||||||      :::|||      |||   :::   |||
Db        465 AATCGTGTCATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTC 524

Qy        150 MetGly---GlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHis 168
              ||||||   :::||||||||||   |||:::|||   ||||||||||||||||||||
Db        525 ATGGGGCGTCGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGGATGGCTTTTGCCGC 584

Qy        169 PhePheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeu 188
              ||||||:::   |||   ||||||   |||||||||:::   |||:::||||||||||||
Db        585 TTCTTCAAGAATCCTTTACCGCTCGGCTTCTGGAGAAGATTGCTGATTCAGGCTTTTCTG 644

Qy        189 SerCysLeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeu 203
              |||           ||||||      ::::::|||   ||||||
Db        645 TCAGGCTTCTTTGCAACAGCCATCTTTTTTATCTGGAAACGTTTA 689

RESULT 11
AF067660
LOCUS       AF067660                1225 bp    mRNA    linear   ROD 04-DEC-1998
DEFINITION  Mus musculus Bcl-2 homolog (Diva) mRNA, complete cds.
ACCESSION   AF067660
VERSION     AF067660.1  GI:3955265
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1225)
  AUTHORS   Inohara,N., Gourley,T.S., Carrio,R., Muniz,M., Merino,J.,
            Garcia,I., Koseki,T., Hu,Y., Chen,S. and Nunez,G.
  TITLE     Diva, a Bcl-2 homologue that binds directly to Apaf-1 and induces
            BH3-independent cell death
  JOURNAL   J. Biol. Chem. 273 (49), 32479-32486 (1998)
  MEDLINE   99047617
   PUBMED   9829980
REFERENCE   2  (bases 1 to 1225)
  AUTHORS   Inohara,N. and Nunez,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-1998) Department of Pathology, University of
            Michigan Medical School, 1500 E.Medical Dr., Ann Arbor, MI 48109,
            USA
FEATURES             Location/Qualifiers
     source          1. .1225
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="B6D2F1/J"
                     /db_xref="taxon:10090"
     gene            1. .1225
                     /gene="Diva"
     CDS             143. .718
                     /gene="Diva"
                     /note="DIVA"
                     /codon_start=1
                     /product="Bcl-2 homolog"
                     /protein_id="AAC83150.1"
                     /db_xref="GI:3955266"
                     /translation="MADSQDPLHERTRRLLSDYIFFCAREPDTPEPPPTSVEAALLRS
                     VTRQIQQEHQEFFSSFCESRGNRLELVKQMADKLLSKDQDFSWSQLVMLLAFAGTLMN
                     QGPYMAVKQKRDLGNRVIVTRDCCLIVNFLYNLLMGRRHRARLEALGGWDGFCRFFKN
                     PLPLGFWRRLLIQAFLSGFFATAIFFIWKRL"
     polyA_signal    1197. .1202
                     /gene="Diva"
     polyA_site      1225
                     /gene="Diva"
ORIGIN

Alignment Scores:
Pred. No.:                3.56e-30        Length:          1225
Score:                    474.50          Matches:         94
Percent Similarity:       64.62%          Conservative:    32
Best Local Similarity:    48.21%          Mismatches:      56
Query Match:              43.77%          Indels:          13
DB:                       10              Gaps:            3

US-10-071-174-2 (1-204) x AF067660 (1-1225)

Qy         13 AspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArg 32
              |||||||||   |||||||||       ||||||:::||||||:::   :::|||||||||
Db        158 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 217

Qy         33 GluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArgSerAla 52
              ||||||   ||||||||||||   |||::::::   |||||||||:::||||||||||
Db        218 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 277

Qy         53 AlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyrProGly 72
                  :::::::::|||   |||:::   |||||||||::::::             |||
Db        278 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 337

Qy         73 AsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerProGlyPro 92
              ||||||   |||||||||      ||||||||   :::||||||
Db        338 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 397
```

```
Qy          93 ThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArgGlyPro 112
               :::|||   :::::::|||   |||:::   |||||||||||||||:::   :::|||||||
Db         398 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 457

Qy         113 LeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGly 132
                  :::      :::   |||:::                                |||
Db         458 TACATGGCTGTCAAGCAGAAGAGG-------------------------GATCTGGGG 490

Qy         133 Asp---------ValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeu 149
               :::         |||   ||||||||     :::|||       |||   :::   |||
Db         491 AATCGTGTCATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTC 550

Qy         150 MetGly---GlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHis 168
               ||||||   :::|||||||||   |||:::|||   ||||||||||||||||||||
Db         551 ATGGGGCGTCGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGGATGGCTTTTGCCGC 610

Qy         169 PhePheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeu 188
               ||||||:::   |||   ||||||   |||||||||:::   |||:::|||||||||||||
Db         611 TTCTTCAAGAATCCTTTACCGCTCGGCTTCTGGAGAAGATTGCTGATTCAGGCTTTTCTG 670

Qy         189 SerCysLeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeu 203
               |||             ||||||      ::::::|||   ||||||
Db         671 TCAGGCTTCTTTGCAACAGCCATCTTTTTTATCTGGAAACGTTTA 715

RESULT 12
BC052690
LOCUS       BC052690                1257 bp    mRNA    linear   ROD 08-OCT-2003
DEFINITION  Mus musculus Bcl2-like 10, mRNA (cDNA clone MGC:60542
            IMAGE:30052580), complete cds.
ACCESSION   BC052690
VERSION     BC052690.1  GI:30851238
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1257)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1257)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Minoru Ko
            cDNA Library Preparation: Yulan Piao and Minoru Ko (National
            Institute on Aging, NIH: http://lgsun.grc.nia.nih.gov/cDNA/)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc_mgc@nhgri.nih.gov
            Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
            Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
            Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
            Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
            Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
            Young,A., Zhang,L.-H. and Green,E.D.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 112 Row: h Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 7304926.
FEATURES             Location/Qualifiers
     source          1. .1257
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="MGC:60542 IMAGE:30052580"
                     /tissue_type="Egg, unfertilized, mouse"
                     /clone_lib="NIA Mouse Unfertilized Egg cDNA Library
                     (Long)"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1"
     gene            1. .1257
                     /gene="Bcl2l10"
                     /note="synonyms: Boo, Diva"
                     /db_xref="LocusID:12049"
                     /db_xref="MGI:1330841"
     CDS             145. .720
                     /gene="Bcl2l10"
                     /codon_start=1
                     /product="Bcl2l10 protein"
                     /protein_id="AAH52690.1"
                     /db_xref="GI:30851239"
                     /db_xref="LocusID:12049"
                     /db_xref="MGI:1330841"
                     /translation="MADSQDPLHGRTRRLLSDYIFFCAREPDTPEPPPTSVEAALLRS
                     VTRQIQQEHQEFFSSFCESRGNRLELVKQMADKLLSKDQDFSWSQLVMLLAFAGTLMN
                     QGPYMAVKQKRDLGNRVIVTRDCCLIVNFLYNLLMGRRHRARLEALGGWDGFCRFFKN
                     PLPLGFWRRLLIQAFLSGFFATAIFFIWKRL"
     misc_feature    268. .597
                     /gene="Bcl2l10"
                     /note="BCL; Region: BCL (B-Cell lymphoma)"
                     /db_xref="CDD:smart00337"
ORIGIN

Alignment Scores:
Pred. No.:                1.39e-29        Length:          1257
Score:                    467.50          Matches:         93
Percent Similarity:       64.10%          Conservative:    32
Best Local Similarity:    47.69%          Mismatches:      57
Query Match:              43.13%          Indels:          13
DB:                       10              Gaps:            3

US-10-071-174-2 (1-204) x BC052690 (1-1257)

Qy          13 AspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArg 32
               ||||||||      ||||||      ||||||:::||||||:::   :::|||||||||
Db         160 GACCCACTGCATGGACGCACTAGACGGCTGCTGTCTCTGACTACATATTCTTCTGCGCACGG 219

Qy          33 GluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArgSerAla 52
               ||||||   ||||||||||||   |||::::::   ||||||||||:::||||||||||
Db         220 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 279
```

```
Qy          53 AlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyrProGly 72
Db         280 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 339

Qy          73 AsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerProGlyPro 92
Db         340 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 399

Qy          93 ThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArgGlyPro 112
Db         400 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 459

Qy         113 LeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGly 132
Db         460 TACATGGCTGTCAAGCAGAAGAGG-------------------------GATCTGGGG 492

Qy         133 Asp---------ValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeu 149
Db         493 AATCGTGTCATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTC 552

Qy         150 MetGly---GlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHis 168
Db         553 ATGGGGCGTCGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGGATGGCTTTTGCCGC 612

Qy         169 PhePheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeu 188
Db         613 TTCTTCAAGAATCCTTTACCGCTCGGCTTCTGGAGAAGATTGCTGATTCAGGCTTTTCTG 672

Qy         189 SerCysLeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeu 203
Db         673 TCAGGCTTCTTTGCAACAGCCATCTTTTTTATCTGGAAACGTTTA 717
```

RESULT 13
AC111669

```
LOCUS       AC111669               240461 bp    DNA     linear   HTG 09-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-54I10, WORKING DRAFT SEQUENCE, 5
            unordered pieces.
ACCESSION   AC111669
VERSION     AC111669.4  GI:24818904
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 240461)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 240461)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 240461)
  AUTHORS   Rat Genome Sequencing Consortium.,
  TITLE     Direct Submission
  JOURNAL   Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 9, 2002 this sequence version replaced gi:23196052.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GMVB
                Center clone name: CH230-54I10
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 218898 bases at least Q40
                Consensus quality: 222128 bases at least Q30
                Consensus quality: 223928 bases at least Q20
                Estimated insert size: 223140; sum-of-contigs estimation
                Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
```

```
            * as soon as it is available and the accession number will
            * be preserved.
            *         1     9792: contig of 9792 bp in length
            *      9793     9892: gap of unknown length
            *      9893   118024: contig of 108132 bp in length
            *    118025   118124: gap of unknown length
            *    118125   236958: contig of 118834 bp in length
            *    236959   237058: gap of unknown length
            *    237059   238398: contig of 1340 bp in length
            *    238399   238498: gap of unknown length
            *    238499   240461: contig of 1963 bp in length.
FEATURES             Location/Qualifiers
     source          1. .240461
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-54I10"
     misc_feature    1. .1351
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    6797. .7581
                     /note="clone_boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH349014"
     misc_feature    8640. .9792
                     /note="wgs_contig"
     misc_feature    118125. .119133
                     /note="wgs_contig"
ORIGIN

Alignment Scores:
Pred. No.:              9.11e-17        Length:         240461
Score:                  342.00          Matches:        69
Percent Similarity:     63.46%          Conservative:   30
Best Local Similarity:  44.23%          Mismatches:     44
Query Match:            31.55%          Indels:         13
DB:                     2               Gaps:           2

US-10-071-174-2 (1-204) x AC111669 (1-240461)

Qy         11 MetAlaAspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCys 30
              |||   |||||||||:::::::||||||      ||||||   ||||||:::   :::|||
Db      21810 ATGGGTGACCCGCTGCAGGATCGCACTAGACGGCTGCTGACTGACTACATATTGTTCTGC 21869

Qy         31 AlaArgGluProGlyThrProGluProAlaProSerThrProGluAla-AlaValLeuAr 50
              ||||||   |||   ||||||||||||   |||::::::   ||||||| |||:::|||||
Db      21870 GCACGGGCGCCGAACACCCCTGAGCCACTGCCCACGTCTGTTGAGGCGCGCCTTGCTGCG 21929

Qy         50 gSerAlaAlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTy 70
              ||||      ::::::::::::|||   |||:::      |||:::::::::      ||
Db      21930 CTCTGTGACTAGTCAGATCCAACAGGAGCACCAGGATCTTTTCAACTCCTTCCGCGACTA 21989

Qy         70 rProGlyAsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerPr 90
              |   |||||||||   |||||||||      |||||||||   :::|||||:::
Db      21990 CCAGGGCAACCGCCTGGAGCTGGTGACACAGATGGCGGATGAGTTGCTCTCCAATGACCA 22049

Qy         90 oGlyProThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluAr 110
                  ||||||||:::|||   |||:::   |||   ||||||||:::   ::
Db      22050 AGAGTTCAACTGGGGCCGCCTGGTGATGCTCCTGGCCTTCGTGGGGACGCTAATGAACCA 22109

Qy        110 gGlyProLeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGl 130
              :      |||      |||                          :::::::||
Db      22110 AGACAGGACTGTTAAGCGGAGG----------------------------AGGGATCA 22139

Qy        130 nGluGlyAsp------ValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerAr 148
              |          :::   |||||||||      :::|||:::||||||   :::||
Db      22140 AAGAAACCGTCTCCTACTGGAGCGAGACTGCTATCTCATAGTGAGCTTGCTGTACAATCG 22199

Qy        148 gLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrp 163
              ||||   |||:::||||||:::||||||:::|||   |||||||||
Db      22200 ACTCACAGGACGGCATCGCTCCTGGCTGGAGGCTCACGGTGGCTGG 22245
```

```
RESULT 14
AC133947
LOCUS       AC133947              169914 bp    DNA     linear   HTG 10-JUL-2004
DEFINITION  Mus musculus chromosome 9 clone RP24-189I2, WORKING DRAFT SEQUENCE,
            3 unordered pieces.
ACCESSION   AC133947
VERSION     AC133947.2  GI:50199129
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 169914)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 169914)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-SEP-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 169914)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Jul 10, 2004 this sequence version replaced gi:23238076.

            ---------------- Genome Center ----------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information -------------------
            Center project name: M_BB0189I02
            ------------------------------------------------------
            -------------- Summary Statistics  -------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 167792 bases at least Q40
            Consensus quality: 168148 bases at least Q30
            Consensus quality: 168291 bases at least Q20
            ------------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    21204: contig of 21204 bp in length
            *    21205    21304: gap of unknown length
            *    21305    86045: contig of 64741 bp in length
            *    86046    86145: gap of unknown length
            *    86146   169914: contig of 83769 bp in length.
FEATURES             Location/Qualifiers
     source          1. .169914
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /clone="RP24-189I2"
     misc_feature    1. .21204
                     /note="assembly_name:Contig13"
     misc_feature    21305. .86045
                     /note="assembly_name:Contig14"
     misc_feature    86146. .169914
                     /note="assembly_name:Contig15
                     clone_end:SP6
```

vector_side:right"
ORIGIN

Alignment Scores:

Pred. No.:	3.21e-16	Length:	169914
Score:	333.50	Matches:	71
Percent Similarity:	63.23%	Conservative:	27
Best Local Similarity:	45.81%	Mismatches:	44
Query Match:	30.77%	Indels:	13
DB:	2	Gaps:	3

US-10-071-174-2 (1-204) x AC133947 (1-169914)

```
Qy         13 AspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArg 32
              ||||||||   |||||||||       ||||||:::|||||:::   :::|||||||||
Db      83475 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 83534

Qy         33 GluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArgSerAla 52
              ||||||  ||||||||||||   |||::::::   ||||||||:::|||||||||
Db      83535 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 83594

Qy         53 AlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyrProGly 72
                  :::::::::|||   |||:::   ||||||||||::::::             |||
Db      83595 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 83654

Qy         73 AsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerProGlyPro 92
              ||||||   |||||||||       |||||||||   :::||||||
Db      83655 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 83714

Qy         93 ThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArgGlyPro 112
              :::|||   :::::::|||   |||:::   |||||||||||||||:::   :::||||||
Db      83715 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 83774

Qy        113 LeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGly 132
                 :::   :::  |||:::                                    |||
Db      83775 TACATGGCTGTCAAGCAGAAGAGG-------------------------GATCTGGGG 83807

Qy        133 Asp---------ValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeu 149
              :::      |||   |||||||||      :::|||      |||   :::   |||
Db      83808 AATCGTGTCATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTC 83867

Qy        150 MetGly---GlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrp 163
              ||||||   :::||||||||||   |||:::|||   |||||||||
Db      83868 ATGGGGCGTCGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGG 83912
```

RESULT 15
AC115880

```
LOCUS       AC115880               202851 bp    DNA     linear   ROD 02-APR-2004
DEFINITION  Mus musculus chromosome 9, clone RP24-365N15, complete sequence.
ACCESSION   AC115880
VERSION     AC115880.11  GI:46063808
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 202851)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 9, clone RP24-365N15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 202851)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 202851)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 202851)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 2, 2004 this sequence version replaced gi:44681569.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@broad.mit.edu
            -------------- Project Information
                Center project name: L24773
```

```
                Center clone name: 365_N_15
             --------------.
FEATURES             Location/Qualifiers
     source          1. .202851
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /map="9"
                     /clone="RP24-365N15"
                     /clone_lib="RPCI-24 Male Mouse BAC"
     misc_feature    1. .2331
                     /note="wgs_end_extension
                     clone_end:SP6"
     repeat_region   350. .517
                     /rpt_family="CYRA11_MM"
     repeat_region   405. .517
                     /rpt_family="URR1B"
     repeat_region   complement(518. .845)
                     /rpt_family="ORR1A2"
     repeat_region   complement(898. .1107)
                     /rpt_family="MLT1D"
     repeat_region   complement(2325. .2432)
                     /rpt_family="ID_B1"
     misc_feature    2332. .2335
                     /note="clone_boundary
                     clone_end:SP6
                     site:MboI"
     repeat_region   3266. .3401
                     /rpt_family="B1_MM"
     repeat_region   3440. .3626
                     /rpt_family="B2_Mm2"
     repeat_region   complement(3934. .4068)
                     /rpt_family="B2_Mm2"
     repeat_region   5256. .5305
                     /rpt_family="(TTTC)n"
     repeat_region   5400. .5555
                     /rpt_family="B3"
     repeat_region   5556. .5642
                     /rpt_family="(CATA)n"
     repeat_region   5643. .5684
                     /rpt_family="B3"
     repeat_region   6066. .6101
                     /rpt_family="(CACAG)n"
     repeat_region   6783. .6968
                     /rpt_family="(TGAG)n"
     repeat_region   7058. .7106
                     /rpt_family="(GA)n"
     repeat_region   7139. .7283
                     /rpt_family="B1_MM"
     repeat_region   7287. .7326
                     /rpt_family="(GAAA)n"
     repeat_region   complement(7627. .7727)
                     /rpt_family="B1F"
     repeat_region   complement(7650. .7758)
                     /rpt_family="B4"
     repeat_region   8301. .8421
                     /rpt_family="PB1D10"
     repeat_region   8422. .8449
                     /rpt_family="ID_B1"
     repeat_region   9184. .9220
                     /rpt_family="(TCTCC)n"
     repeat_region   complement(10225. .10356)
                     /rpt_family="B1_MM"
     repeat_region   10495. .10572
                     /rpt_family="ID1_MM"
     repeat_region   complement(10673. .10947)
                     /rpt_family="MLT1F"
     repeat_region   complement(10978. .11186)
                     /rpt_family="MTD"
     repeat_region   complement(11187. .11580)
                     /rpt_family="ORR1A-int"
     repeat_region   11582. .11707
                     /rpt_family="B1F"
     repeat_region   11776. .11824
                     /rpt_family="(CA)n"
     repeat_region   11913. .12060
                     /rpt_family="B3A"
     repeat_region   12061. .12103
                     /rpt_family="(CA)n"
     repeat_region   12104. .12134
                     /rpt_family="B3A"
     repeat_region   complement(12393. .12493)
                     /rpt_family="PB1D7"
     repeat_region   complement(12620. .12631)
                     /rpt_family="ID5"
     repeat_region   complement(12632. .12757)
                     /rpt_family="B1F"
     repeat_region   complement(12758. .12829)
                     /rpt_family="ID_B1"
     repeat_region   complement(12758. .12821)
                     /rpt_family="ID5"
     repeat_region   complement(12956. .13319)
                     /rpt_family="ORR1B2"
     repeat_region   13921. .14047
                     /rpt_family="B1F"
     repeat_region   15511. .15586
                     /rpt_family="B4A"
     repeat_region   15591. .15722
                     /rpt_family="RSINE1"
     repeat_region   15724. .15775

Alignment Scores:
Pred. No.:              3.87e-16        Length:         202851
Score:                  333.50          Matches:        71
Percent Similarity:     63.23%          Conservative:   27
Best Local Similarity:  45.81%          Mismatches:     44
Query Match:            30.77%          Indels:         13
DB:                     10              Gaps:           3

US-10-071-174-2 (1-204) x AC115880 (1-202851)

Qy         13 AspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArg 32
              |||||||||   |||||||||       ||||||:::||||||:::   :::|||||||||
Db      30352 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 30411

Qy         33 GluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArgSerAla 52
              ||||||   ||||||||||||   |||:::::::   |||||||||:::|||||||||
Db      30412 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 30471

Qy         53 AlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyrProGly 72
                  :::::::::|||   |||:::   |||||||||::::::             |||
Db      30472 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 30531

Qy         73 AsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerProGlyPro 92
              ||||||   |||||||||       |||||||||   :::||||||
Db      30532 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 30591

Qy         93 ThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArgGlyPro 112
              :::|||   ::::::|||   |||:::   |||||||||||||||:::   :::|||||||
Db      30592 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 30651

Qy        113 LeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGly 132
                 :::      :::   |||:::                                |||
Db      30652 TACATGGCTGTCAAGCAGAAGAGG-------------------------GATCTGGGG 30684

Qy        133 Asp---------ValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeu 149
                 :::         |||   |||||||||       :::|||      |||   :::   |||
Db      30685 AATCGTGTCATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTC 30744

Qy        150 MetGly---GlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrp 163
              ||||||   :::|||||||||   |||:::|||   |||||||||
Db      30745 ATGGGGCGTCGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGG 30789


Search completed: November  7, 2004, 07:45:29
```

Job time : 4754 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 6, 2004, 21:41:06 ; Search time 6234 Seconds
(without alignments)
6728.579 Million cell updates/sec

Title: US-10-071-174-1
Perfect score: 887
Sequence: 1 cgggccaagaaaaccagcga.....ctcttccttgagtgaaagaa 887

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		887	100.0	887	9	AF326964	AF326964 Homo sapi
2		812.4	91.6	1168	6	BD233466	BD233466 Human pro
3		615	69.3	615	9	AF285092	AF285092 Homo sapi
4		585	66.0	585	9	HSA458330	AJ458330 Homo sapi
5		582	65.6	582	6	BD233456	BD233456 Human pro
6		574.4	64.8	726	6	CQ752105	CQ752105 Sequence
7	c	539.2	60.8	93287	9	AC023906	AC023906 Homo sapi
8	c	345.6	39.0	214669	2	AC018903	AC018903 Homo sapi
9		327.6	36.9	214669	2	AC018903	AC018903 Homo sapi
10		259.2	29.2	1209	10	AF102501	AF102501 Mus muscu
11		259.2	29.2	1225	10	AF067660	AF067660 Mus muscu
12		257.6	29.0	1257	10	BC052690	BC052690 Mus muscu
13		244.2	27.5	1074	10	AY029163	AY029163 Rattus no
14		147	16.6	169914	2	AC133947	AC133947 Mus muscu
15		147	16.6	202851	10	AC115880	AC115880 Mus muscu
16		141.8	16.0	240461	2	AC111669	AC111669 Rattus no
17		95.4	10.8	243686	2	AC096430	AC096430 Rattus no
18	c	85.2	9.6	160755	9	AC104976	AC104976 Homo sapi
19		68	7.7	2000	6	AX655393	AX655393 Sequence
20	c	62.2	7.0	295150	1	SCO939126	AL939126 Streptomy
21	c	61.2	6.9	125020	9	AF429315	AF429315 Homo sapi
22		60	6.8	60	6	CQ545914	CQ545914 Sequence
23		58.2	6.6	9471	5	AY186728	AY186728 Gallus ga
24		58.2	6.6	125020	9	AF429315	AF429315 Homo sapi
25	c	57.4	6.5	82746	1	AF453501	AF453501 Actinosyn
26		57.2	6.4	3987	9	BC036533	BC036533 Homo sapi
27	c	57.2	6.4	300100	1	SCO939123	AL939123 Streptomy
28		57	6.4	1822	5	GGCHOX14	X52670 Chicken Cho
29		56.8	6.4	66669	1	AME16952	Y16952 Amycolatops
30		56.8	6.4	314100	1	SCO939106	AL939106 Streptomy
31	c	56.4	6.4	963	11	PM12H12G	AL684456 Penicilli
32		55.6	6.3	943	9	AB042637S3	AB042639 Homo sapi
33		55.6	6.3	2507	9	AB042636	AB042636 Homo sapi
34		55.6	6.3	2811	9	AK055486	AK055486 Homo sapi
35		55.6	6.3	4023	9	AK126663	AK126663 Homo sapi
36	c	55.6	6.3	97110	2	AC016815	AC016815 Homo sapi
37	c	55.6	6.3	137960	2	AC103889	AC103889 Homo sapi
38		55.6	6.3	168133	2	AC093622	AC093622 Homo sapi
39	c	55.6	6.3	208236	9	AC010536	AC010536 Homo sapi
40		55.2	6.2	1265	8	AK108845	AK108845 Oryza sat
41		55.2	6.2	2997	6	AX552616	AX552616 Sequence
42		55.2	6.2	3158	6	AX552614	AX552614 Sequence
43	c	55.2	6.2	99575	2	AP004588	AP004588 Oryza sat
44	c	55.2	6.2	140142	8	AP005871	AP005871 Oryza sat
45	c	55.2	6.2	144879	8	AP005405	AP005405 Oryza sat

ALIGNMENTS

RESULT 1
AF326964
LOCUS AF326964 887 bp mRNA linear PRI 01-MAY-2001
DEFINITION Homo sapiens BCLB (BCLB) mRNA, complete cds.
ACCESSION AF326964
VERSION AF326964.1 GI:13898393
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS Ke,N., Godzik,A. and Reed,J.C.
TITLE Bcl-B, a novel Bcl-2 family member that differentially binds and regulates Bax and Bak
JOURNAL J. Biol. Chem. 276 (16), 12481-12484 (2001)
MEDLINE 21201065
PUBMED 11278245
REFERENCE 2 (bases 1 to 887)
AUTHORS Ke,N., Godzik,A. and Reed,J.C.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2000) The Burnham Institute, 10901 N. Torrey Pines Rd., La Jolla, CA 92037, USA
FEATURES Location/Qualifiers
source 1. .887
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q21"
gene 1. .887
/gene="BCLB"
CDS 50. .664
/gene="BCLB"
/note="BCL2 family protein; Boo-like"
/codon_start=1
/product="BCLB"
/protein_id="AAK48715.1"
/db_xref="GI:13898394"
/translation="MVDQLRERTTMADPLRERTELLLADYLGYCAREPGTPEPAPSTP
EAAVLRSAAARLRQIHRSFFSAYLGYPGNRFELVALMADSVLSDSPGPTWGRVVTLVT
FAGTLLERGPLVTARWKKWGFQPRLKEQEGDVARDCQRLVALLSSRLMGQHRAWLQAQ

GGWDGFCHFFRTPFPLAFWRKQLVQAFLSCLLTTAFIYLWTRLL"

ORIGIN

Query Match 100.0%; Score 887; DB 9; Length 887;
Best Local Similarity 100.0%; Pred. No. 5.4e-144;
Matches 887; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 60

Qy         61 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 120

Qy        121 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 180

Qy        181 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 240

Qy        241 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 300

Qy        301 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 360

Qy        361 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 420

Qy        421 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 480

Qy        481 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGA 540

Qy        541 TGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGT 600

Qy        601 CCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATT 660

Qy        661 ATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAG 720

Qy        721 ATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGA 780

Qy        781 ATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACT 840

Qy        841 CAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
```

RESULT 2
BD233466
LOCUS BD233466 1168 bp DNA linear PAT 17-JUL-2003
DEFINITION Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION BD233466
VERSION BD233466.1 GI:33043236
KEYWORDS JP 2002519016-A/12.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1168)
AUTHORS Kato,S. and Kimura,T.
TITLE Human protein having hydrophobic domain and DNA encoding the same
JOURNAL Patent: JP 2002519016-A 12 02-JUL-2002;
SAGAMI CHEMICAL RESEARCH CENTER,PROTEGENE INC
COMMENT OS Homo sapiens (human)
PN JP 2002519016-A/12
PD 02-JUL-2002
PF 18-JUN-1999 JP 2000557267
PI SEISHI KATO,TOMOKO KIMURA
PC
C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
00
CC Human protein having hydrophobic domain and DNA encoding the
CC same
FH Key Location/Qualifiers
FT source 1. .1168
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1. .1168
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 91.6%; Score 812.4; DB 6; Length 1168;
Best Local Similarity 99.9%; Pred. No. 4.5e-131;
Matches 813; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         74 ACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGG 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGG 60

Qy        134 TACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTG 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTG 120

Qy        194 CTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTC 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTC 180

Qy        254 GGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGAC 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGAC 240

Qy        314 AGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTG 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTG 300

Qy        374 GAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAG 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAG 360

Qy        434 GAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGG 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGG 420

Qy        494 CTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCAC 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCAC 480

Qy        554 TTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTG 613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTG 540

Qy        614 TCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAA 673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        541 TCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAA 600

Qy        674 CTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACA 733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACA 660

Qy        734 AGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGA 793
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        661 AGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTACCTGAATGCATACAAGGA 720

Qy        794 GTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGAC 853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGAC 780

Qy        854 AAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
              |||||||||||||||||||||||||||||||||
Db        781 AAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 814


RESULT 3
AF285092
LOCUS       AF285092                615 bp    mRNA    linear   PRI 08-NOV-2001
DEFINITION  Homo sapiens Bcl-2-like protein 10 mRNA, complete cds.
ACCESSION   AF285092
VERSION     AF285092.1  GI:9837265
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 615)
  AUTHORS   Zhang,H., Holzgreve,W. and De Geyter,C.
  TITLE     Bcl2-L-10, a novel anti-apoptotic member of the Bcl-2 family,
            blocks apoptosis in the mitochondria death pathway but not in the
            death receptor pathway
  JOURNAL   Hum. Mol. Genet. 10 (21), 2329-2339 (2001)
  MEDLINE   21548034
   PUBMED   11689480
REFERENCE   2  (bases 1 to 615)
  AUTHORS   Zhang,H.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2000) University Women's Hospital, Schanzenstra
            46, Basel 4057, Switzerland
FEATURES             Location/Qualifiers
     source          1. .615
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21"
                     /tissue_type="ovary"
     CDS             1. .615
                     /codon_start=1
                     /product="Bcl-2-like protein 10"
                     /protein_id="AAG00503.1"
                     /db_xref="GI:9837266"
                     /translation="MVDQLRERTTMADPLRERTELLLADYLGYCAREPGTPEPAPSTP
                     EAAVLRSAAARLRQIHRSFFSAYLGYPGNRFELVALMADSVLSDSPGPTWGRVVTLVT
                     FAGTLLERGPLVTARWKKWGFQPRLKEQEGDVARDCQRLVALLSSRLMGQHRAWLQAQ
                     GGWDGFCHFFRTPFPLAFWRKQLVQAFLSCLLTTAFIYLWTRLL"
ORIGIN

  Query Match           69.3%;  Score 615;  DB 9;  Length 615;
  Best Local Similarity  100.0%;  Pred. No. 8e-97;
  Matches  615;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         50 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 60

Qy        110 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 120

Qy        170 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 180

Qy        230 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 240

Qy        290 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 300

Qy        350 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 360

Qy        410 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 420

Qy        470 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 480

Qy        530 GGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGA 589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGA 540

Qy        590 AAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGG 649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGG 600

Qy        650 ACACGATTATTATGA 664
              |||||||||||||||
Db        601 ACACGATTATTATGA 615


RESULT 4
HSA458330
LOCUS       HSA458330               585 bp    DNA     linear   PRI 25-APR-2002
DEFINITION  Homo sapiens NRH gene for anti-apoptotic protein.
ACCESSION   AJ458330
VERSION     AJ458330.1  GI:20338765
KEYWORDS    anti-apoptotic protein; NRH gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Aouacheria,A., Arnaud,E., Venet,S., Lalle,P., Gouy,M., Rigal,D. and
            Gillet,G.
  TITLE     Nrh, a human homologue of Nr-13 associates with Bcl-Xs and is an
            inhibitor of apoptosis
  JOURNAL   Oncogene 20 (41), 5846-5855 (2001)
  MEDLINE   21477277
   PUBMED   11593390
REFERENCE   2  (bases 1 to 585)
  AUTHORS   Gillet,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-2002) Gillet G., Ibcp, CNRS UMR 5086, 7 PASSAGE
            DU VERCORS, 69367, FRANCE
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
     gene            1. .585
                     /gene="NRH"
     CDS             1. .585
                     /gene="NRH"
                     /function="apoptosis inhibitor"
```

```
                     /codon_start=1
                     /product="anti-apoptotic protein"
                     /protein_id="CAD30221.1"
                     /db_xref="GI:20338766"
                     /db_xref="GOA:Q9HD36"
                     /db_xref="Swiss-Prot:Q9HD36"
                     /translation="MADPLRERTELLLADYLGYCAREPGTPEPAPSTPEAAVLRSAAA
                     RLRQIHRSFFSAYLGYPGNRFELVALMADSVLSDSPGPTWGRVVTLVTFAGTLLERGP
                     LVTARWKKWGFQPRLKEQEGDVARDCQRLVALLSSRLMGQHRAWLQAQGGWDGFCHFF
                     RTPFPLAFWRKQLVQAFLSCLLTTAFIYLWTRLL"
ORIGIN

  Query Match             66.0%;  Score 585;  DB 9;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-91;
  Matches  585;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         80 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 60

Qy        140 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 199
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 120

Qy        200 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 259
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 180

Qy        260 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 319
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 240

Qy        320 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 379
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 300

Qy        380 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 439
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 360

Qy        440 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 499
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 420

Qy        500 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 559
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 480

Qy        560 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 619
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 540

Qy        620 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGA 664
               |||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGA 585


RESULT 5
BD233456
LOCUS       BD233456                582 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION   BD233456
VERSION     BD233456.1  GI:33043226
KEYWORDS    JP 2002519016-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Kato,S. and Kimura,T.
  TITLE     Human protein having hydrophobic domain and DNA encoding the same
  JOURNAL   Patent: JP 2002519016-A 2 02-JUL-2002;
            SAGAMI CHEMICAL RESEARCH CENTER,PROTEGENE INC
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002519016-A/2
            PD   02-JUL-2002
            PF   18-JUN-1999 JP 2000557267
            PI   SEISHI KATO,TOMOKO KIMURA
            PC
            C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
             00
            CC   Human protein having hydrophobic domain and DNA encoding the
            CC             same
            FH   Key             Location/Qualifiers
            FT   source          1. .582
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             65.6%;  Score 582;  DB 6;  Length 582;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-91;
  Matches  582;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         80 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 60

Qy        140 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 199
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 120

Qy        200 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 259
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 180

Qy        260 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 319
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 240

Qy        320 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 379
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 300

Qy        380 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 439
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 360

Qy        440 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 499
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 420

Qy        500 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 559
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 480

Qy        560 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 619
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 540

Qy        620 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 661
               |||||||||||||||||||||||||||||||||||||||||
Db        541 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 582


RESULT 6
CQ752105
LOCUS       CQ752105                726 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 38039 from Patent WO02068579.
ACCESSION   CQ752105
VERSION     CQ752105.1  GI:42387450
KEYWORDS    .
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 38039 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             64.8%;  Score 574.4;  DB 6;  Length 726;
  Best Local Similarity   90.2%;  Pred. No. 8.6e-90;
  Matches  655;  Conservative    0;  Mismatches    1;  Indels   70;  Gaps    1;

Qy         50 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 60

Qy        110 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 169
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 120

Qy        170 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 180

Qy        230 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 240

Qy        290 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 300

Qy        350 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 360

Qy        410 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 420

Qy        470 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 480

Qy        530 GGCGGCTG---------------------------------------------------- 537
              ||||||||
Db        481 GGCGGCTGGGTGAGCACGCGGCGGACACCGGGACACGGGGCGGGACGGGCAGCCGGGAAG 540

Qy        538 ------------------GGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGC 579
                               ||||||||||||||||||||||||||||||||||||||||||
Db        541 CGCCCACGAGGCTGGCACGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGC 600

Qy        580 TTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCAT 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCAT 660

Qy        640 TTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAA 699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAA 720

Qy        700 CTGTGA 705
              ||||||
Db        721 CTGTGA 726
```

```
RESULT 7
AC023906/c
LOCUS       AC023906               93287 bp    DNA     linear   PRI 06-JUL-2001
DEFINITION  Homo sapiens chromosome 15 clone CTD-2184D3 map 15q21.2, complete
            sequence.
ACCESSION   AC023906
VERSION     AC023906.7  GI:14595770
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2000) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
REFERENCE   4  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
COMMENT     On Jul 4, 2001 this sequence version replaced gi:12248292.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399

            Note:  Data from overlapping BACs AC010674 [drafting center:
            UWMSC], AC090970 [drafting center: UWMSC]; and AC016824 [drafting
            center: GTC] was added for finishing
            ------------.
FEATURES             Location/Qualifiers
     source          1. .93287
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21.2"
                     /clone="CTD-2184D3"
                     /clone_lib="Cal Tech Human BAC library D"
                     /note="Data from overlapping clones CTD-2650P22 AC090970,
                     RP11-47K1 AC016824, and RP11-430B1 was added and the
                     consensus was determined from CTD-2184D3 to the extent
                     possible"
     misc_feature    1. .8893
```

```
                     /note="overlap with CTD-2650P22 AC090970"
     misc_feature    1. .8893
                     /note="overlap with RP11-47K1 AC016824"
     unsure          6078. .6140
                     /note="low quality data"
     unsure          10529. .10532
                     /note="low quality data"
     unsure          14722. .14808
                     /note="low quality data"
     unsure          37627
                     /note="low quality data"
     unsure          71552. .71593
                     /note="low quality data"
     misc_feature    82012. .93287
                     /note="overlap with RP11-430B1 AC010674"
ORIGIN

  Query Match             60.8%;  Score 539.2;  DB 9;  Length 93287;
  Best Local Similarity   99.4%;  Pred. No. 7.5e-84;
  Matches  541;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64991 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 64932

Qy         61 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64931 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 64872

Qy        121 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64871 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 64812

Qy        181 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64811 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 64752

Qy        241 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64751 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 64692

Qy        301 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64691 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 64632

Qy        361 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64631 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 64572

Qy        421 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64571 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 64512

Qy        481 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64511 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGT 64452

Qy        541 TGGC 544
                ||
Db      64451 GAGC 64448


RESULT 8
AC018903/c
LOCUS       AC018903               214669 bp    DNA     linear   HTG 04-JUN-2000
DEFINITION  Homo sapiens chromosome 15 clone RP11-337B11 map 15q21, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC018903
VERSION     AC018903.2  GI:8247797
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 214669)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., Kaur,A.,
            Madan,A., Nesbitt,R., Shaffer,T. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 214669)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-1999) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
COMMENT     On Jun 4, 2000 this sequence version replaced gi:6630517.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399
            ------------.
            * NOTE: This record contains 192 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      806: contig of 806 bp in length
            *      807      906: gap of unknown length
            *      907     2068: contig of 1162 bp in length
            *     2069     2168: gap of unknown length
            *     2169     3020: contig of 852 bp in length
            *     3021     3120: gap of unknown length
            *     3121     4284: contig of 1164 bp in length
            *     4285     4384: gap of unknown length
            *     4385     5201: contig of 817 bp in length
            *     5202     5301: gap of unknown length
            *     5302     6544: contig of 1243 bp in length
            *     6545     6644: gap of unknown length
            *     6645     7472: contig of 828 bp in length
            *     7473     7572: gap of unknown length
            *     7573     8756: contig of 1184 bp in length
            *     8757     8856: gap of unknown length
            *     8857     9658: contig of 802 bp in length
            *     9659     9758: gap of unknown length
            *     9759    10905: contig of 1147 bp in length
            *    10906    11005: gap of unknown length
            *    11006    11811: contig of 806 bp in length
            *    11812    11911: gap of unknown length
            *    11912    13293: contig of 1382 bp in length
            *    13294    13393: gap of unknown length
            *    13394    14211: contig of 818 bp in length
            *    14212    14311: gap of unknown length
            *    14312    15465: contig of 1154 bp in length
            *    15466    15565: gap of unknown length
            *    15566    16379: contig of 814 bp in length
            *    16380    16479: gap of unknown length
            *    16480    17624: contig of 1145 bp in length
            *    17625    17724: gap of unknown length
            *    17725    18578: contig of 854 bp in length
            *    18579    18678: gap of unknown length
            *    18679    19774: contig of 1096 bp in length
```

```
*    19775    19874: gap of unknown length
*    19875    20700: contig of 826 bp in length
*    20701    20800: gap of unknown length
*    20801    22036: contig of 1236 bp in length
*    22037    22136: gap of unknown length
*    22137    23023: contig of 887 bp in length
*    23024    23123: gap of unknown length
*    23124    24196: contig of 1073 bp in length
*    24197    24296: gap of unknown length
*    24297    25141: contig of 845 bp in length
*    25142    25241: gap of unknown length
*    25242    26385: contig of 1144 bp in length
*    26386    26485: gap of unknown length
*    26486    27325: contig of 840 bp in length
*    27326    27425: gap of unknown length
*    27426    28691: contig of 1266 bp in length
*    28692    28791: gap of unknown length
*    28792    29655: contig of 864 bp in length
*    29656    29755: gap of unknown length
*    29756    30881: contig of 1126 bp in length
*    30882    30981: gap of unknown length
*    30982    31813: contig of 832 bp in length
*    31814    31913: gap of unknown length
*    31914    33056: contig of 1143 bp in length
*    33057    33156: gap of unknown length
*    33157    34007: contig of 851 bp in length
*    34008    34107: gap of unknown length
*    34108    35320: contig of 1213 bp in length
*    35321    35420: gap of unknown length
*    35421    36207: contig of 787 bp in length
*    36208    36307: gap of unknown length
*    36308    37521: contig of 1214 bp in length
*    37522    37621: gap of unknown length
*    37622    38420: contig of 799 bp in length
*    38421    38520: gap of unknown length
*    38521    39726: contig of 1206 bp in length
*    39727    39826: gap of unknown length
*    39827    40635: contig of 809 bp in length
*    40636    40735: gap of unknown length
*    40736    41845: contig of 1110 bp in length
*    41846    41945: gap of unknown length
*    41946    42730: contig of 785 bp in length
*    42731    42830: gap of unknown length
*    42831    43917: contig of 1087 bp in length
*    43918    44017: gap of unknown length
*    44018    44825: contig of 808 bp in length
*    44826    44925: gap of unknown length
*    44926    46114: contig of 1189 bp in length
*    46115    46214: gap of unknown length
*    46215    47031: contig of 817 bp in length
*    47032    47131: gap of unknown length
*    47132    48292: contig of 1161 bp in length
*    48293    48392: gap of unknown length
*    48393    49176: contig of 784 bp in length
*    49177    49276: gap of unknown length
*    49277    50389: contig of 1113 bp in length
*    50390    50489: gap of unknown length
*    50490    51333: contig of 844 bp in length
*    51334    51433: gap of unknown length
*    51434    52765: contig of 1332 bp in length
*    52766    52865: gap of unknown length
*    52866    53710: contig of 845 bp in length
*    53711    53810: gap of unknown length
*    53811    54965: contig of 1155 bp in length
*    54966    55065: gap of unknown length
*    55066    55910: contig of 845 bp in length
*    55911    56010: gap of unknown length
*    56011    57558: contig of 1548 bp in length
*    57559    57658: gap of unknown length
*    57659    58509: contig of 851 bp in length
*    58510    58609: gap of unknown length
*    58610    59745: contig of 1136 bp in length
*    59746    59845: gap of unknown length
*    59846    60655: contig of 810 bp in length
*    60656    60755: gap of unknown length
*    60756    61976: contig of 1221 bp in length
*    61977    62076: gap of unknown length
*    62077    62896: contig of 820 bp in length
*    62897    62996: gap of unknown length
*    62997    64128: contig of 1132 bp in length
*    64129    64228: gap of unknown length
*    64229    65026: contig of 798 bp in length
*    65027    65126: gap of unknown length
*    65127    66269: contig of 1143 bp in length
*    66270    66369: gap of unknown length
*    66370    67220: contig of 851 bp in length
*    67221    67320: gap of unknown length
*    67321    68425: contig of 1105 bp in length
*    68426    68525: gap of unknown length
*    68526    69322: contig of 797 bp in length
*    69323    69422: gap of unknown length
*    69423    70568: contig of 1146 bp in length
*    70569    70668: gap of unknown length
*    70669    71547: contig of 879 bp in length
*    71548    71647: gap of unknown length
*    71648    73629: contig of 1982 bp in length
*    73630    73729: gap of unknown length
*    73730    74587: contig of 858 bp in length
*    74588    74687: gap of unknown length
*    74688    75780: contig of 1093 bp in length
*    75781    75880: gap of unknown length
*    75881    77196: contig of 1316 bp in length
*    77197    77296: gap of unknown length
*    77297    78651: contig of 1355 bp in length
*    78652    78751: gap of unknown length
*    78752    79557: contig of 806 bp in length
*    79558    79657: gap of unknown length
*    79658    80774: contig of 1117 bp in length
*    80775    80874: gap of unknown length
*    80875    81690: contig of 816 bp in length
*    81691    81790: gap of unknown length
*    81791    82928: contig of 1138 bp in length
*    82929    83028: gap of unknown length
*    83029    83902: contig of 874 bp in length
*    83903    84002: gap of unknown length
*    84003    85193: contig of 1191 bp in length
*    85194    85293: gap of unknown length
*    85294    86149: contig of 856 bp in length

Query Match            39.0%;  Score 345.6;  DB 2;  Length 214669;
Best Local Similarity  98.9%;  Pred. No. 2.6e-50;
Matches  348;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        536 TGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAG 595
              | |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     168582 TAGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGAAGAAAACAG 168523

Qy        596 CTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGA 655
              ||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||||
Db     168522 CTGGTCCAGGCTTTTTTGTCATGCCTGTTAACAACAGCCTTCATTTATCTCTGGACACGA 168463

Qy        656 TTATTATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAAT 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168462 TTATTATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAAT 168403

Qy        716 GACAGATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTA 775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168402 GACAGATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTA 168343

Qy        776 TCTGAATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAA 835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168342 TCTGAATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAA 168283

Qy        836 GTACTCAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168282 GTACTCAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 168231
```

```
RESULT 9
AC018903
LOCUS       AC018903               214669 bp    DNA     linear   HTG 04-JUN-2000
DEFINITION  Homo sapiens chromosome 15 clone RP11-337B11 map 15q21, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC018903
VERSION     AC018903.2  GI:8247797
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 214669)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., Kaur,A.,
            Madan,A., Nesbitt,R., Shaffer,T. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 214669)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-1999) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
COMMENT     On Jun 4, 2000 this sequence version replaced gi:6630517.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399
            ------------.
            * NOTE: This record contains 192 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      806: contig of 806 bp in length
            *      807      906: gap of unknown length
            *      907     2068: contig of 1162 bp in length
            *     2069     2168: gap of unknown length
            *     2169     3020: contig of 852 bp in length
            *     3021     3120: gap of unknown length
            *     3121     4284: contig of 1164 bp in length
            *     4285     4384: gap of unknown length
            *     4385     5201: contig of 817 bp in length
            *     5202     5301: gap of unknown length
            *     5302     6544: contig of 1243 bp in length
            *     6545     6644: gap of unknown length
            *     6645     7472: contig of 828 bp in length
            *     7473     7572: gap of unknown length
            *     7573     8756: contig of 1184 bp in length
            *     8757     8856: gap of unknown length
            *     8857     9658: contig of 802 bp in length
            *     9659     9758: gap of unknown length
            *     9759    10905: contig of 1147 bp in length
            *    10906    11005: gap of unknown length
            *    11006    11811: contig of 806 bp in length
            *    11812    11911: gap of unknown length
            *    11912    13293: contig of 1382 bp in length
            *    13294    13393: gap of unknown length
            *    13394    14211: contig of 818 bp in length
            *    14212    14311: gap of unknown length
            *    14312    15465: contig of 1154 bp in length
            *    15466    15565: gap of unknown length
            *    15566    16379: contig of 814 bp in length
            *    16380    16479: gap of unknown length
            *    16480    17624: contig of 1145 bp in length
            *    17625    17724: gap of unknown length
            *    17725    18578: contig of 854 bp in length
            *    18579    18678: gap of unknown length
            *    18679    19774: contig of 1096 bp in length
            *    19775    19874: gap of unknown length
            *    19875    20700: contig of 826 bp in length
            *    20701    20800: gap of unknown length
            *    20801    22036: contig of 1236 bp in length
            *    22037    22136: gap of unknown length
            *    22137    23023: contig of 887 bp in length
            *    23024    23123: gap of unknown length
            *    23124    24196: contig of 1073 bp in length
            *    24197    24296: gap of unknown length
            *    24297    25141: contig of 845 bp in length
            *    25142    25241: gap of unknown length
            *    25242    26385: contig of 1144 bp in length
            *    26386    26485: gap of unknown length
            *    26486    27325: contig of 840 bp in length
            *    27326    27425: gap of unknown length
            *    27426    28691: contig of 1266 bp in length
            *    28692    28791: gap of unknown length
            *    28792    29655: contig of 864 bp in length
            *    29656    29755: gap of unknown length
            *    29756    30881: contig of 1126 bp in length
            *    30882    30981: gap of unknown length
            *    30982    31813: contig of 832 bp in length
            *    31814    31913: gap of unknown length
            *    31914    33056: contig of 1143 bp in length
            *    33057    33156: gap of unknown length
            *    33157    34007: contig of 851 bp in length
            *    34008    34107: gap of unknown length
            *    34108    35320: contig of 1213 bp in length
            *    35321    35420: gap of unknown length
            *    35421    36207: contig of 787 bp in length
            *    36208    36307: gap of unknown length
            *    36308    37521: contig of 1214 bp in length
            *    37522    37621: gap of unknown length
            *    37622    38420: contig of 799 bp in length
            *    38421    38520: gap of unknown length
            *    38521    39726: contig of 1206 bp in length
            *    39727    39826: gap of unknown length
            *    39827    40635: contig of 809 bp in length
            *    40636    40735: gap of unknown length
            *    40736    41845: contig of 1110 bp in length
            *    41846    41945: gap of unknown length
            *    41946    42730: contig of 785 bp in length
            *    42731    42830: gap of unknown length
            *    42831    43917: contig of 1087 bp in length
            *    43918    44017: gap of unknown length
            *    44018    44825: contig of 808 bp in length
            *    44826    44925: gap of unknown length
            *    44926    46114: contig of 1189 bp in length
            *    46115    46214: gap of unknown length
            *    46215    47031: contig of 817 bp in length
            *    47032    47131: gap of unknown length
            *    47132    48292: contig of 1161 bp in length
            *    48293    48392: gap of unknown length
            *    48393    49176: contig of 784 bp in length
            *    49177    49276: gap of unknown length
            *    49277    50389: contig of 1113 bp in length
            *    50390    50489: gap of unknown length
            *    50490    51333: contig of 844 bp in length
            *    51334    51433: gap of unknown length
            *    51434    52765: contig of 1332 bp in length
```

```
            *    52766    52865: gap of unknown length
            *    52866    53710: contig of 845 bp in length
            *    53711    53810: gap of unknown length
            *    53811    54965: contig of 1155 bp in length
            *    54966    55065: gap of unknown length
            *    55066    55910: contig of 845 bp in length
            *    55911    56010: gap of unknown length
            *    56011    57558: contig of 1548 bp in length
            *    57559    57658: gap of unknown length
            *    57659    58509: contig of 851 bp in length
            *    58510    58609: gap of unknown length
            *    58610    59745: contig of 1136 bp in length
            *    59746    59845: gap of unknown length
            *    59846    60655: contig of 810 bp in length
            *    60656    60755: gap of unknown length
            *    60756    61976: contig of 1221 bp in length
            *    61977    62076: gap of unknown length
            *    62077    62896: contig of 820 bp in length
            *    62897    62996: gap of unknown length
            *    62997    64128: contig of 1132 bp in length
            *    64129    64228: gap of unknown length
            *    64229    65026: contig of 798 bp in length
            *    65027    65126: gap of unknown length
            *    65127    66269: contig of 1143 bp in length
            *    66270    66369: gap of unknown length
            *    66370    67220: contig of 851 bp in length
            *    67221    67320: gap of unknown length
            *    67321    68425: contig of 1105 bp in length
            *    68426    68525: gap of unknown length
            *    68526    69322: contig of 797 bp in length
            *    69323    69422: gap of unknown length
            *    69423    70568: contig of 1146 bp in length
            *    70569    70668: gap of unknown length
            *    70669    71547: contig of 879 bp in length
            *    71548    71647: gap of unknown length
            *    71648    73629: contig of 1982 bp in length
            *    73630    73729: gap of unknown length
            *    73730    74587: contig of 858 bp in length
            *    74588    74687: gap of unknown length
            *    74688    75780: contig of 1093 bp in length
            *    75781    75880: gap of unknown length
            *    75881    77196: contig of 1316 bp in length
            *    77197    77296: gap of unknown length
            *    77297    78651: contig of 1355 bp in length
            *    78652    78751: gap of unknown length
            *    78752    79557: contig of 806 bp in length
            *    79558    79657: gap of unknown length
            *    79658    80774: contig of 1117 bp in length
            *    80775    80874: gap of unknown length
            *    80875    81690: contig of 816 bp in length
            *    81691    81790: gap of unknown length
            *    81791    82928: contig of 1138 bp in length
            *    82929    83028: gap of unknown length
            *    83029    83902: contig of 874 bp in length
            *    83903    84002: gap of unknown length
            *    84003    85193: contig of 1191 bp in length
            *    85194    85293: gap of unknown length
            *    85294    86149: contig of 856 bp in length

  Query Match             36.9%;  Score 327.6;  DB 2;  Length 214669;
  Best Local Similarity   80.9%;  Pred. No. 3.4e-47;
  Matches  406;  Conservative    0;  Mismatches   94;  Indels    2;  Gaps    2;

Qy          1 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 60
              | || | |  |||    ||||| || | ||||||||||||||||||||||||||||||
Db      93213 CCGGGCCAAGAAATTTTCGAAGGGCCCGGCCCCCAGCAGAGGCCGGACCATGGTTGACCA 93272

Qy         61 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 120
               ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db      93273 TTTGCGGGAGCGCACCACCATGTGCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGCG 93332

Qy        121 CGACTACCTGGGGTA-CTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGC 179
              |||||||||  | ||| |||| ||||||||||  |||||||||||||  ||||| ||| |
Db      93333 CGACTACCTAGAGTACCTGCCCCCGGGAACCTCGCACCCCCGAGCCGAGGCCATTCACAC 93392

Qy        180 CCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTT 239
              |||||||||||| |||  |||||||| |||  ||||||||||||||| ||||||||||||
Db      93393 CCGAGGCCGCCGAGCTTGGCTCCGCAGCCCGCAGGTTACGGCAGAATCACCGGTCCTTTT 93452

Qy        240 TCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATT 299
              |   | |||||||||| |||||||  |||||||||| ||||||||||||||||||||||
Db      93453 TTCTCCCCTACCTCGGATACCCCGAAAACCGCTTCGACCTGGTGGCGCTGATGGCGGATT 93512

Qy        300 CCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCG 359
              |||| |||||||||||| ||||  ||||||| |||||  |  || ||||||||||||||||
Db      93513 CCGTACTCTCCGACAGACCCGTGCCCACCTAGGGCAAAAAGGGGACGCTCGTGACCTTCT 93572

Qy        360 CAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCC 419
              ||||||| || || |||| | |||| || |||||| ||| | ||| ||     |||| |
Db      93573 CAGGGACACTTCTTGAGAAAAGGCCCCTAGTGACCCCCCCCTTGAAAAAAAATAGCTTTC 93632

Qy        420 AGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCT 479
              | || ||| |||| || ||  || ||  |  |||||  |||||||| | |||||  ||||
Db      93633 AACCCCGGGTAAAAGAACAAAAGAGCACCTGCGCCC-AGACTGCCAACACCTGGAAGCCT 93691

Qy        480 TGCTGAGCTCGCGGCTCATGGG 501
              | |||| |||    || || ||
Db      93692 TACTGAACTCAGAACTTATAGG 93713
```

```
RESULT 10
AF102501
LOCUS       AF102501               1209 bp    mRNA    linear   ROD 21-JAN-1999
DEFINITION  Mus musculus Bcl-2 homolog (Boo) mRNA, complete cds.
ACCESSION   AF102501
VERSION     AF102501.1  GI:4165137
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1209)
  AUTHORS   Song,Q., Kuang,Y., Dixit,V.M. and Vincenz,C.
  TITLE     Boo, a novel negative regulator of cell death, interacts with
            Apaf-1
  JOURNAL   EMBO J. 18 (1), 167-178 (1999)
  MEDLINE   99094902
   PUBMED   9878060
REFERENCE   2  (bases 1 to 1209)
  AUTHORS   Song,Q.Z., Kuang,Y.P., Dixit,V.M. and Vincenz,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-OCT-1998) Pathology, University of Michigan, 1301
            Catherine Road, Ann Arbor, MI 48109, USA
FEATURES             Location/Qualifiers
     source          1. .1209
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="B6D2 F1/J"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /map="D9"
                     /tissue_type="ovary"
     gene            1. .1209
                     /gene="Boo"
     CDS             117. .692
                     /gene="Boo"
                     /codon_start=1
                     /product="Bcl-2 homolog"
                     /protein_id="AAD08703.1"
                     /db_xref="GI:4165138"
                     /translation="MADSQDPLHERTRRLLSDYIFFCAREPDTPEPPPTSVEAALLRS
                     VTRQIQQEHQEFFSSFCESRGNRLELVKQMADKLLSKDQDFSWSQLVMLLAFAGTLMN
                     QGPYMAVKQKRDLGNRVIVTRDCCLIVNFLYNLLMGRRHRARLEALGGWDGFCRFFKN
                     PLPLGFWRRLLIQAFLSGFFATAIFFIWKRL"
ORIGIN
```

```
  Query Match           29.2%;  Score 259.2;  DB 10;  Length 1209;
  Best Local Similarity   63.1%;  Pred. No. 3.7e-35;
  Matches  477;  Conservative    0;  Mismatches  253;  Indels   26;  Gaps    4;

Qy         86 GACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGG 145
Db        132 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 191

Qy        146 GAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCG 205
Db        192 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 251

Qy        206 GCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGG 265
Db        252 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 311

Qy        266 AACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCC 325
Db        312 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 371

Qy        326 ACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCG 385
Db        372 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 431

Qy        386 CTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGC 445
Db        432 TACATGGCTGTCAAGCAGAAGA----------------GGGATCTGGGGAATCGTGTC 473

Qy        446 GACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGG--- 502
Db        474 ATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTCATGGGGCGT 533

Qy        503 CAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGG 562
Db        534 CGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGGATGGCTTTTGCCGCTTCTTCAAG 593

Qy        563 ACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTG 622
Db        594 AATCCTTTACCGCTCGGCTTCTGGAGAAGATTGCTGATTCAGGCTTTTCTGTCAGGCTTC 653

Qy        623 TTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACC 682
Db        654 TTTGCAACAGCCATCTTTTTTATCTGGAAACGTTTAT----AAGTTTAAAATTTTTAAAG 709

Qy        683 CGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAG 742
Db        710 CGATCTTACCTACCTACCTGTGAACCCTCTAAGGAACAATTCTGGGTAATTGGATAAGAA 769

Qy        743 GGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGAGTCCTGAGG 802
Db        770 CTGAGGAAGTCCTCCAGCCTAGAGACATTTCTACCTGCATGC-TACATGGAGTTCTGGGT 828

Qy        803 TGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTA 838
Db        829 TTGTGATGGGGCCAGCATTTGAAAGGGGCCCATGTA 864


RESULT 11
AF067660
LOCUS       AF067660               1225 bp    mRNA    linear   ROD 04-DEC-1998
DEFINITION  Mus musculus Bcl-2 homolog (Diva) mRNA, complete cds.
ACCESSION   AF067660
VERSION     AF067660.1  GI:3955265
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1225)
  AUTHORS   Inohara,N., Gourley,T.S., Carrio,R., Muniz,M., Merino,J.,
            Garcia,I., Koseki,T., Hu,Y., Chen,S. and Nunez,G.
  TITLE     Diva, a Bcl-2 homologue that binds directly to Apaf-1 and induces
            BH3-independent cell death
  JOURNAL   J. Biol. Chem. 273 (49), 32479-32486 (1998)
  MEDLINE   99047617
   PUBMED   9829980
REFERENCE   2  (bases 1 to 1225)
  AUTHORS   Inohara,N. and Nunez,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-MAY-1998) Department of Pathology, University of
            Michigan Medical School, 1500 E.Medical Dr., Ann Arbor, MI 48109,
            USA
FEATURES             Location/Qualifiers
     source          1. .1225
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="B6D2F1/J"
                     /db_xref="taxon:10090"
     gene            1. .1225
                     /gene="Diva"
     CDS             143. .718
                     /gene="Diva"
                     /note="DIVA"
                     /codon_start=1
                     /product="Bcl-2 homolog"
                     /protein_id="AAC83150.1"
                     /db_xref="GI:3955266"
                     /translation="MADSQDPLHERTRRLLSDYIFFCAREPDTPEPPPTSVEAALLRS
                     VTRQIQQEHQEFFSSFCESRGNRLELVKQMADKLLSKDQDFSWSQLVMLLAFAGTLMN
                     QGPYMAVKQKRDLGNRVIVTRDCCLIVNFLYNLLMGRRHRARLEALGGWDGFCRFFKN
                     PLPLGFWRRLLIQAFLSGFFATAIFFIWKRL"
     polyA_signal    1197. .1202
                     /gene="Diva"
     polyA_site      1225
                     /gene="Diva"
ORIGIN

  Query Match           29.2%;  Score 259.2;  DB 10;  Length 1225;
  Best Local Similarity   63.1%;  Pred. No. 3.7e-35;
  Matches  477;  Conservative    0;  Mismatches  253;  Indels   26;  Gaps    4;

Qy         86 GACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGG 145
Db        158 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 217

Qy        146 GAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCG 205
Db        218 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 277

Qy        206 GCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGG 265
Db        278 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 337

Qy        266 AACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCC 325
Db        338 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 397

Qy        326 ACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCG 385
Db        398 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 457

Qy        386 CTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGC 445
Db        458 TACATGGCTGTCAAGCAGAAGA----------------GGGATCTGGGGAATCGTGTC 499

Qy        446 GACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGG--- 502
Db        500 ATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTCATGGGGCGT 559

Qy        503 CAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGG 562
Db        560 CGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGGATGGCTTTTGCCGCTTCTTCAAG 619

Qy        563 ACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTG 622
Db        620 AATCCTTTACCGCTCGGCTTCTGGAGAAGATTGCTGATTCAGGCTTTTCTGTCAGGCTTC 679
```

```
Qy        623 TTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACC 682
              ||  |||||||| || ||| | |||||| ||| ||||     | |||||| ||||||
Db        680 TTTGCAACAGCCATCTTTTTTATCTGGAAACGTTTAT----AAGTTTAAAATTTTTAAAG 735

Qy        683 CGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAG 742
              || |  |||| || | |||||| |      || ||   ||    |  ||   ||   ||
Db        736 CGATCTTACCTACCTACCTGTGAACCCTCTAAGGAACAATTCTGGGTAATTGGATAAGAA 795

Qy        743 GGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGAGTCCTGAGG 802
                | | |     ||   |   |||| ||| || ||| |||| |||| ||||| ||| |
Db        796 CTGAGGAAGTCCTCCAGCCTAGAGACATTTCTACCTGCATGC-TACATGGAGTTCTGGGT 854

Qy        803 TGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTA 838
              | |||||  ||||||  ||| ||   |   || |||
Db        855 TTGTGATGGGGCCAGCATTTGAAAGGGGCCCATGTA 890
```

RESULT 12
BC052690
LOCUS BC052690 1257 bp mRNA linear ROD 08-OCT-2003
DEFINITION Mus musculus Bcl2-like 10, mRNA (cDNA clone MGC:60542 IMAGE:30052580), complete cds.
ACCESSION BC052690
VERSION BC052690.1 GI:30851238
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1257)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 1257)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Minoru Ko
cDNA Library Preparation: Yulan Piao and Minoru Ko (National Institute on Aging, NIH: http://lgsun.grc.nia.nih.gov/cDNA/)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 112 Row: h Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7304926.

FEATURES Location/Qualifiers
source 1. .1257
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="MGC:60542 IMAGE:30052580"
/tissue_type="Egg, unfertilized, mouse"
/clone_lib="NIA Mouse Unfertilized Egg cDNA Library (Long)"
/lab_host="DH10B"
/note="Vector: pSPORT1"
gene 1. .1257
/gene="Bcl2l10"
/note="synonyms: Boo, Diva"
/db_xref="LocusID:12049"
/db_xref="MGI:1330841"
CDS 145. .720
/gene="Bcl2l10"
/codon_start=1
/product="Bcl2l10 protein"
/protein_id="AAH52690.1"
/db_xref="GI:30851239"
/db_xref="LocusID:12049"
/db_xref="MGI:1330841"
/translation="MADSQDPLHGRTRRLLSDYIFFCAREPDTPEPPPTSVEAALLRS VTRQIQQEHQEFFSSFCESRGNRLELVKQMADKLLSKDQDFSWSQLVMLLAFAGTLMN QGPYMAVKQRDLGNRVIVTRDCCLIVNFLYNLLMGRRHRARLEALGGWDGFCRFFKN PLPLGFWRRLLIQAFLSGFFATAIFFIWKRL"
misc_feature 268. .597
/gene="Bcl2l10"
/note="BCL; Region: BCL (B-Cell lymphoma)"
/db_xref="CDD:smart00337"

ORIGIN

Query Match 29.0%; Score 257.6; DB 10; Length 1257;
Best Local Similarity 63.0%; Pred. No. 6.9e-35;
Matches 476; Conservative 0; Mismatches 254; Indels 26; Gaps 4;

```
Qy         86 GACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGG 145
              ||||| ||||  |  |||||     | | ||||| | |||||| |     | ||||||  |||
Db        160 GACCCACTGCATGGACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 219

Qy        146 GAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCG 205
              || || | |||||| |||||  ||||  |  |   |||||| ||| |||| ||||| | |
Db        220 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 279

Qy        206 GCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGG 265
               |     | | | ||||   ||||      ||||| ||| ||| |  ||    ||  ||
Db        280 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 339

Qy        266 AACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCC 325
              || ||| | |||||||||    | |||||| |||     |||||||| |    ||  | | |
Db        340 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 399

Qy        326 ACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCG 385
              | |||| ||  | |||||| |||| || ||||||| ||||||||  ||  |  ||| ||
Db        400 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 459
```

```
Qy        386 CTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGC 445
Db        460 TACATGGCTGTCAAGCAGAAGA----------------GGGATCTGGGGAATCGTGTC 501

Qy        446 GACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGG--- 502
Db        502 ATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTCATGGGGCGT 561

Qy        503 CAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGG 562
Db        562 CGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGGATGGCTTTTGCCGCTTCTTCAAG 621

Qy        563 ACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTG 622
Db        622 AATCCTTTACCGCTCGGCTTCTGGAGAAGATTGCTGATTCAGGCTTTTCTGTCAGGCTTC 681

Qy        623 TTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACC 682
Db        682 TTTGCAACAGCCATCTTTTTTATCTGGAAACGTTTAT----AAGTTTAAAATTTTTAAAG 737

Qy        683 CGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAG 742
Db        738 CGATCTTACCTACCTACCTGTGAACCCTCTAAGGAACAATTCTGGGTAATTGGATAAGAA 797

Qy        743 GGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGAGTCCTGAGG 802
Db        798 CTGAGGAAGTCCTCCAGCCTAGAGACATTTCTACCTGCATGC-TACATGGAGTTCTGGGT 856

Qy        803 TGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTA 838
Db        857 TTGTGATGGGGCCAGCATTTGAAAGGGGCCCATGTA 892
```

RESULT 13
AY029163
LOCUS AY029163 1074 bp mRNA linear ROD 19-JUN-2003
DEFINITION Rattus norvegicus BCL2L10 (Bcl2l10) mRNA, complete cds.
ACCESSION AY029163
VERSION AY029163.1 GI:13641257
KEYWORDS .
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 1074)
AUTHORS Itoh,T., Itoh,A. and Pleasure,D.
TITLE Bcl-2-related protein family gene expression during oligodendroglial differentiation
JOURNAL J. Neurochem. 85 (6), 1500-1512 (2003)
MEDLINE 22672518
PUBMED 12787069
REFERENCE 2 (bases 1 to 1074)
AUTHORS Itoh,T., Itoh,A. and Pleasure,D.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2001) Neurology Research, The Children's Hospital of Philadelphia, Abramson Research Center, Room 516 I, 3517 Civic Center Boulevard, Philadelphia, PA 19104, USA
FEATURES Location/Qualifiers
source 1. .1074
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
gene 1. .1074
/gene="Bcl2l10"
CDS 12. .569
/gene="Bcl2l10"
/note="Boo, Diva"
/codon_start=1
/product="BCL2L10"
/protein_id="AAK31792.1"
/db_xref="GI:13641258"
/translation="MGDPLQDRTRRLLTDYILFCARAPNTPEPLPTSVEAALLRSVTS
QIQQEHQDLFNSFRDYQGNRLELVTQMADELLSNDQEFNWGRLVMLLAFVGTLMNQDR
TVKRRRDQRNRLLLERDCYLIVSLLYNRLTGRHRSWLEAHGGWDGFCQFFKNPLPPGF
WRRLLIRAILSCFFATAIFYIWKCL"
ORIGIN

Query Match 27.5%; Score 244.2; DB 10; Length 1074;
Best Local Similarity 61.7%; Pred. No. 1.5e-32;
Matches 455; Conservative 0; Mismatches 253; Indels 29; Gaps 3;

```
Qy         80 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 139
Db         12 ATGGGTGACCCGCTGCAGGATCGCACTAGACGGCTGCTGACTGACTACATATTGTTCTGC 71

Qy        140 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 199
Db         72 GCACGGGCGCCGAACACCCCTGAGCCACTGCCCACGTCTGTTGAGGCGGCCTTGCTGCGC 131

Qy        200 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 259
Db        132 TCTGTGACTAGTCAGATCCAACAGGAGCACCAGGATCTTTTCAACTCCTTCCGCGACTAC 191

Qy        260 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 319
Db        192 CAGGGCAACCGCCTGGAGCTGGTGACACAGATGGCGGATGAGTTGCTCTCCAATGACCAA 251

Qy        320 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 379
Db        252 GAGTTCAACTGGGGCCGCCTGGTGATGCTCCTGGCCTTCGTGGGGACGCTAATGAACCAA 311

Qy        380 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 439
Db        312 GACAGGACTGTTAAGCGGAGGAGGGATCAAAGAAACCGTCTCCTAC-------------- 357

Qy        440 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 499
Db        358 ----------TGGAGCGAGACTGCTATCTCATAGTGAGCTTGCTGTACAATCGACTCACA 407

Qy        500 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 559
Db        408 GGACGGCATCGCTCCTGGCTGGAGGCTCACGGTGGCTGGGATGGCTTTTGCCAATTCTTC 467

Qy        560 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 619
Db        468 AAGAACCCCTTACCACCCGGCTTCTGGAGAAGATTGCTGATCCGGGCTATTCTGTCCTGT 527

Qy        620 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTA 679
Db        528 TTCTTTGCAACGGCCATCTTTTATATCTGGAAATGTTTAT----AAGTTTTAAATTTTTA 583

Qy        680 ACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACT 739
Db        584 AATCGATCTTACCTGCCTAACTGCGGGCCCTCTAAGGGACAATTCTGGGTAGTTGGATAA 643

Qy        740 GAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGAGTCCTG 799
Db        644 GAACTGAGGAAATCCTCCTGCCTAGAGACATTTCTACCTGCATGC-TACACGGAGTTCCG 702

Qy        800 AGGTGGTGATTTGGCCA 816
Db        703 GGTTTGGGATGGGCCTA 719
```

RESULT 14
AC133947
LOCUS AC133947 169914 bp DNA linear HTG 10-JUL-2004
DEFINITION Mus musculus chromosome 9 clone RP24-189I2, WORKING DRAFT SEQUENCE, 3 unordered pieces.
ACCESSION AC133947
VERSION AC133947.2 GI:50199129
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 169914)
  AUTHORS   Wilson,R.K.
  TITLE     The sequence of Mus musculus clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 169914)
  AUTHORS   McPherson,J.D. and Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-SEP-2002) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
REFERENCE   3  (bases 1 to 169914)
  AUTHORS   Wilson,R.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2004) Genome Sequencing Center, 4444 Forest Park
            Parkway, St. Louis, MO 63108, USA
COMMENT     On Jul 10, 2004 this sequence version replaced gi:23238076.

            ----------------- Genome Center ----------------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu
            Contact: submissions@watson.wustl.edu
            -------------- Project Information -------------------
            Center project name: M_BB0189I02
            ------------------------------------------------------
            -------------- Summary Statistics  -------------
            Sequencing vector: M13; 0%
            Sequencing vector: plasmid; 100%
            Chemistry: Dye-primer ET; 0% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 167792 bases at least Q40
            Consensus quality: 168148 bases at least Q30
            Consensus quality: 168291 bases at least Q20
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    21204: contig of 21204 bp in length
            *    21205    21304: gap of unknown length
            *    21305    86045: contig of 64741 bp in length
            *    86046    86145: gap of unknown length
            *    86146   169914: contig of 83769 bp in length.
FEATURES             Location/Qualifiers
     source          1. .169914
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /clone="RP24-189I2"
     misc_feature    1. .21204
                     /note="assembly_name:Contig13"
     misc_feature    21305. .86045
                     /note="assembly_name:Contig14"
     misc_feature    86146. .169914
                     /note="assembly_name:Contig15
                     clone_end:SP6
                     vector_side:right"
ORIGIN

  Query Match             16.6%;  Score 147;  DB 2;  Length 169914;
  Best Local Similarity   61.5%;  Pred. No. 7e-16;
  Matches  281;  Conservative    0;  Mismatches  155;  Indels   21;  Gaps    2;

Qy         86 GACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGG 145
              ||||| ||||  || |||||    | | ||||| | |||||| |    | ||||||
Db      83475 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 83534

Qy        146 GAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCG 205
              || || | |||||| |||||  ||||  |  |    |||||| ||| |||| ||||| | |
Db      83535 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 83594

Qy        206 GCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGG 265
              |      | | | ||||   ||||     ||||| ||| ||| |  ||    ||  ||
Db      83595 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 83654

Qy        266 AACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCC 325
              || ||| | |||||||||   | |||||| |||     |||||||| |   ||  | | |
Db      83655 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 83714

Qy        326 ACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCG 385
              | |||| ||  | |||||| |||| || |||||||| ||||||||  || |   ||| ||
Db      83715 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 83774

Qy        386 CTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGC 445
                || | | |  |||||                    ||     || | |   | |
Db      83775 TACATGGCTGTCAAGCAGAAGA----------------GGGATCTGGGGAATCGTGTC 83816

Qy        446 GACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGG--- 502
              ||  |||| ||||||   | | | |||  ||| |||      | ||||||||||
Db      83817 ATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTCATGGGGCGT 83876

Qy        503 CAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGG 539
              | |||||||||| ||||| ||||||  ||||||||||
Db      83877 CGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGG 83913


RESULT 15
AC115880
LOCUS       AC115880               202851 bp    DNA     linear   ROD 02-APR-2004
DEFINITION  Mus musculus chromosome 9, clone RP24-365N15, complete sequence.
ACCESSION   AC115880
VERSION     AC115880.11  GI:46063808
KEYWORDS    HTG.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 202851)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Mus musculus chromosome 9, clone RP24-365N15
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 202851)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
            Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
            Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 202851)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
```

```
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 202851)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
            Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
            Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
            DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
            Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
            Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
            Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
            Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
            MacLean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
            Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
            Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 2, 2004 this sequence version replaced gi:44681569.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@broad.mit.edu
            -------------- Project Information
                Center project name: L24773
                Center clone name: 365_N_15
            --------------.
FEATURES             Location/Qualifiers
     source          1. .202851
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="9"
                     /map="9"
                     /clone="RP24-365N15"
                     /clone_lib="RPCI-24 Male Mouse BAC"
     misc_feature    1. .2331
                     /note="wgs_end_extension
                     clone_end:SP6"
     repeat_region   350. .517
                     /rpt_family="CYRA11_MM"
     repeat_region   405. .517
                     /rpt_family="URR1B"
     repeat_region   complement(518. .845)
                     /rpt_family="ORR1A2"
     repeat_region   complement(898. .1107)
                     /rpt_family="MLT1D"
     repeat_region   complement(2325. .2432)
                     /rpt_family="ID_B1"
     misc_feature    2332. .2335
                     /note="clone_boundary
                     clone_end:SP6
                     site:MboI"
     repeat_region   3266. .3401
                     /rpt_family="B1_MM"
     repeat_region   3440. .3626
                     /rpt_family="B2_Mm2"
     repeat_region   complement(3934. .4068)
                     /rpt_family="B2_Mm2"
     repeat_region   5256. .5305
                     /rpt_family="(TTTC)n"
     repeat_region   5400. .5555
                     /rpt_family="B3"
     repeat_region   5556. .5642
                     /rpt_family="(CATA)n"
     repeat_region   5643. .5684
                     /rpt_family="B3"
     repeat_region   6066. .6101
                     /rpt_family="(CACAG)n"
     repeat_region   6783. .6968
                     /rpt_family="(TGAG)n"
     repeat_region   7058. .7106
                     /rpt_family="(GA)n"
     repeat_region   7139. .7283
                     /rpt_family="B1_MM"
     repeat_region   7287. .7326
                     /rpt_family="(GAAA)n"
     repeat_region   complement(7627. .7727)
                     /rpt_family="B1F"
     repeat_region   complement(7650. .7758)
                     /rpt_family="B4"
     repeat_region   8301. .8421
                     /rpt_family="PB1D10"
     repeat_region   8422. .8449
                     /rpt_family="ID_B1"
     repeat_region   9184. .9220
                     /rpt_family="(TCTCC)n"
     repeat_region   complement(10225. .10356)
                     /rpt_family="B1_MM"
     repeat_region   10495. .10572
                     /rpt_family="ID1_MM"
     repeat_region   complement(10673. .10947)
                     /rpt_family="MLT1F"
     repeat_region   complement(10978. .11186)
                     /rpt_family="MTD"
     repeat_region   complement(11187. .11580)
                     /rpt_family="ORR1A-int"
     repeat_region   11582. .11707
                     /rpt_family="B1F"
     repeat_region   11776. .11824
                     /rpt_family="(CA)n"
     repeat_region   11913. .12060
                     /rpt_family="B3A"
     repeat_region   12061. .12103
                     /rpt_family="(CA)n"
     repeat_region   12104. .12134
                     /rpt_family="B3A"
     repeat_region   complement(12393. .12493)
                     /rpt_family="PB1D7"
     repeat_region   complement(12620. .12631)
                     /rpt_family="ID5"
     repeat_region   complement(12632. .12757)
                     /rpt_family="B1F"
```

```
     repeat_region   complement(12758. .12829)
                     /rpt_family="ID_B1"
     repeat_region   complement(12758. .12821)
                     /rpt_family="ID5"
     repeat_region   complement(12956. .13319)
                     /rpt_family="ORR1B2"
     repeat_region   13921. .14047
                     /rpt_family="B1F"
     repeat_region   15511. .15586
                     /rpt_family="B4A"
     repeat_region   15591. .15722
                     /rpt_family="RSINE1"
     repeat_region   15724. .15775

  Query Match             16.6%;  Score 147;  DB 10;  Length 202851;
  Best Local Similarity   61.5%;  Pred. No. 6.9e-16;
  Matches  281;  Conservative    0;  Mismatches  155;  Indels   21;  Gaps    2;

Qy          86 GACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGG 145
               ||||| |||| || |||||    | | ||||| | |||||| |    | ||||||| |||
Db       30352 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 30411

Qy         146 GAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCG 205
               || || | |||||| |||||  ||||  |  |   |||||| ||| |||| ||||| | |
Db       30412 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 30471

Qy         206 GCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGG 265
                 |     | | | ||||   ||||     |||||| ||| ||| |  ||    ||  ||
Db       30472 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 30531

Qy         266 AACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCC 325
               || ||| | |||||||||   | |||||| |||    |||||||| |   ||  | | |
Db       30532 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 30591

Qy         326 ACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCG 385
               | |||| ||  | |||||| |||| || ||||||| ||||||||  || |   ||| ||
Db       30592 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 30651

Qy         386 CTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGC 445
                 || | | |  |  |||||                    ||     || | |    | |
Db       30652 TACATGGCTGTCAAGCAGAAGA----------------GGGATCTGGGGAATCGTGTC 30693

Qy         446 GACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGG--- 502
                ||  |||| |||||   | | | |||  ||| |||      | ||||||||||
Db       30694 ATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTCATGGGGCGT 30753

Qy         503 CAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGG 539
               | |||||||||| ||||| ||||||  ||||||||||
Db       30754 CGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGG 30790


Search completed: November  7, 2004, 01:20:03
Job time : 6240 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 7, 2004, 03:03:17 ; Search time 510 Seconds
(without alignments)
2099.768 Million cell updates/sec

Title: US-10-071-174-2
Perfect score: 1084
Sequence: 1 MVDQLRERTTMADPLRERTE.....QAFLSCLLTTAFIYLWTRLL 204

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10071174/runat_05112004_131903_23974/app_query.fasta_1.391
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10071174_@CGN_1_1_885_@runat_05112004_131903_23974 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1084	100.0	887	6	AAD46683	Aad46683 Human Bcl
2	1045	96.4	1168	3	AAZ90049	Aaz90049 Hydrophob
3	1035	95.5	582	3	AAZ90039	Aaz90039 Hydrophob
4	701	64.7	522	4	AAH47022	Aah47022 Human Bcl
5	639	58.9	548	4	ABK41913	Abk41913 cDNA enco
6	639	58.9	548	9	ADB59580	Adb59580 Connectiv
7	559	51.6	874	6	ABQ44403	Abq44403 Oligonucl
c 8	559	51.6	874	6	ABQ44402	Abq44402 Oligonucl
9	513	47.3	874	6	ABQ44404	Abq44404 Oligonucl
c 10	513	47.3	874	6	ABQ44405	Abq44405 Oligonucl
11	475.5	43.9	650	10	ADD34132	Add34132 Mouse mit
c 12	232.5	21.4	835	10	ADD34582	Add34582 Mouse mit
13	199	18.4	338	10	ADD34583	Add34583 Mouse mit
14	129	11.9	658	6	ABT09344	Abt09344 Phase-1 R
15	129	11.9	658	10	ADG30848	Adg30848 Liver tox
16	129	11.9	658	12	ADG45405	Adg45405 Liver inf
17	129	11.9	658	12	ADH22707	Adh22707 Partial D
18	127	11.7	573	10	ADH26628	Adh26628 Human Bax
19	127	11.7	579	2	AAZ19764	Aaz19764 Human wil
20	127	11.7	579	4	AAC84598	Aac84598 Human Bax
21	127	11.7	579	4	AAC90811	Aac90811 Human Bax
22	127	11.7	579	6	ABV78149	Abv78149 Human Bax
23	127	11.7	579	6	ABZ35725	Abz35725 Human bax
24	127	11.7	579	6	ABX09968	Abx09968 Human Bax
25	127	11.7	579	6	ABL91690	Abl91690 Human pol
26	127	11.7	579	9	ADA20845	Ada20845 Human BAX
27	127	11.7	579	9	AAL61054	Aal61054 Human Bax
28	127	11.7	579	11	ADI32061	Adi32061 Human cDN
29	127	11.7	579	12	ADO05979	Ado05979 Pro-apopt
30	127	11.7	624	2	AAQ97606	Aaq97606 Human Bax
31	127	11.7	779	2	AAZ15935	Aaz15935 Human gen
32	127	11.7	791	2	AAZ77545	Aaz77545 Human ova
33	127	11.7	996	2	AAZ06732	Aaz06732 Nucleotid
34	126	11.6	677	9	ADA20844	Ada20844 Human BAX
35	125	11.5	624	2	AAV84005	Aav84005 cDNA enco
36	125	11.5	624	4	AAF77704	Aaf77704 Human bcl
37	120	11.1	596	2	ADF39373	Adf39373 Human BAX
38	120	11.1	1607	2	ADF39391	Adf39391 Human BAX
39	119	11.0	579	10	ADB52764	Adb52764 Primary r
40	119	11.0	579	10	ABT41943	Abt41943 Toxicity
41	118	10.9	579	6	AAF88666	Aaf88666 C. albica
42	118	10.9	583	2	AAT96577	Aat96577 Human bcl
43	118	10.9	583	2	AAX25134	Aax25134 Human bcl
44	117.5	10.8	581	2	AAT96578	Aat96578 Mouse bcl
45	117.5	10.8	581	2	AAX25135	Aax25135 Mouse bcl

ALIGNMENTS

RESULT 1
AAD46683
ID AAD46683 standard; DNA; 887 BP.
XX
AC AAD46683;
XX
DT 27-JAN-2003 (first entry)
XX
DE Human Bcl-B DNA.
XX
KW Human; Bcl-2; Bcl-B; therapy; apoptosis; cell degenerative disorder;
KW proliferative disorder; muscle degeneration; Alzheimer's disease; CJD;
KW Creutzfeldt-Jacob's disease; Machado-Joseph disease; MJD; transgenic;
KW Parkinson's disease; Huntington's disease; HD; spinocerebellar ataxia;
KW SCA; dentatorubropallidoluysian atrophy; DRPLA; Kennedy's disease;
KW stroke; ischaemia; head trauma; neoplasia; anticonvulsant; vulnerary;
KW nootropic; neuroprotective; cytostatic; immunosuppressive; vasotropic;
KW cerebroprotective; autoimmune disorder; chromosome 15; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 50. .664
FT /*tag= a
FT /product= "Human Bcl-B protein"
XX
PN WO200272601-A2.
XX
PD 19-SEP-2002.
XX

```
PF   07-FEB-2002; 2002WO-US003547.
XX
PR   07-FEB-2001; 2001US-0267166P.
PR   07-FEB-2002; 2002US-00071174.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Reed JC,  Ke N,  Godzik A;
XX
DR   WPI; 2002-723312/78.
DR   P-PSDB; AAE29097.
XX
PT   New isolated or recombinant Bcl-B nucleic acids and polypeptides, for
PT   treating a disorder associated with apoptosis, such as cell degenerative
PT   or proliferative disorder e.g. cancer, Alzheimer's disease or Parkinson's
PT   disease.
XX
PS   Claim 11; Page 74; 82pp; English.
XX
CC   The invention relates to human member of Bcl-2 family Bcl-B protein and
CC   its corresponding nucleic acid. Bcl-B is useful in treating a subject
CC   having or at risk of a disorder associated with apoptosis, such as cell
CC   degenerative or proliferative disorder like neural or muscle
CC   degeneration, e.g. Alzheimer's disease, Creutzfeldt-Jacob's disease
CC   (CJD), Machado-Joseph disease (MJD), Parkinson's disease, Huntington's
CC   disease (HD), spinocerebellar ataxias 1, 2 and 6 (SCA-1, -2 and -6),
CC   dentatorubropallidoluysian atrophy (DRPLA), Kennedy's disease, stroke,
CC   ischaemia, head trauma, neoplasia, autoimmune disorder or fibrotic
CC   condition. The transgenic animals are used as in vivo models to study
CC   apoptosis and potential therapies for apoptosis. The present sequence is
CC   human Bcl-B DNA. Bcl-B gene is located at chromosome 15
XX
SQ   Sequence 887 BP; 171 A; 270 C; 269 G; 177 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.16e-96      Length:          887
Score:                    1084.00       Matches:         204
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       6             Gaps:            0

US-10-071-174-2 (1-204) x AAD46683 (1-887)

Qy          1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 109

Qy         21 LeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArgGluProGlyThrProGluProAla 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 169

Qy         41 ProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHis 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 229

Qy         61 ArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        230 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 289

Qy         81 MetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLeu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 349

Qy        101 ValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLys 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        350 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 409

Qy        121 TrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArg 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        410 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 469

Qy        141 LeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGln 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        470 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 529

Qy        161 GlyGlyTrpAspGlyPheCysHisPhePheArgThrProPheProLeuAlaPheTrpArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        530 GGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGA 589

Qy        181 LysGlnLeuValGlnAlaPheLeuSerCysLeuLeuThrThrAlaPheIleTyrLeuTrp 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        590 AAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGG 649

Qy        201 ThrArgLeuLeu 204
              ||||||||||||
Db        650 ACACGATTATTA 661

RESULT 2
AAZ90049
ID   AAZ90049 standard; cDNA; 1168 BP.
XX
AC   AAZ90049;
XX
DT   09-MAY-2000  (first entry)
XX
DE   Hydrophobic domain containing protein clone HP02403 nucleotide sequence.
XX
KW   Hydrophobic domain; clone HP02403; nutritional supplement; SCID; HIV;
KW   cell proliferation; immune stimulant; immune deficiency; tumour; pain;
KW   rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
KW   myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
KW   anti-inflammatory; infection; bodily characteristic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200000506-A2.
XX
PD   06-JAN-2000.
XX
PF   18-JUN-1999;   99WO-JP003242.
XX
PR   26-JUN-1998;   98JP-00180008.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Kimura T;
XX
DR   WPI; 2000-160665/14.
DR   P-PSDB; AAY78802.
XX
PT   Novel human proteins having hydrophobic domains used for research and
PT   diagnostic purposes.
XX
PS   Claim 4; Page 92-94; 117pp; English.
XX
CC   This sequence represents the hydrophobic domain containing protein, clone
CC   HP00631 nucleotide sequence. The sequence is isolated from a human
CC   stomach cancer cell line. The HP02403 protein contains one putative
CC   transmembrane domain. The protein shows homology to the japanese quail
CC   apoptosis regualtor NR-13. The invention relates to human proteins with
CC   hydrophobic domains, the DNA and the cDNA encoding them. The
CC   polynucleotides and proteins are predicted to have biological activities
CC   which make them suitable for treating, preventing or ameliorating medical
CC   conditions in humans and animals. Suggested activities include
CC   nutritional activity (nutritional source or supplement); cytokine and
CC   cell proliferation/differentiation activity; immune stimulating (e.g. as
CC   vaccines) or suppressing activity (e.g. to treat various immune
CC   deficiencies such as SCIDS or HIV, connective tissue disease, systemic
CC   lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC   inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC   dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC   and autoimmune inflammatory eye disease, as well as asthma, allergies and
CC   organ transplantation); haematopoiesis regulating activity (e.g. in
CC   treatment of myeloid or lymphoid cell deficiencies); tissue growth
```

CC activity (e.g. wound healing and tissue repair, ulcers, burns,
CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
CC activity; haemostatic and thrombolytic activity (e.g. treating
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and
CC tumour inhibition activity. The polynucleotides are also stated to be
CC useful for gene therapy. Other activities include inhibiting infections
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,
CC skin, effecting biorhythms or caricadic cycles; enhancing fertility;
CC treatment of depression; treatment of pain; hormonal or endocrine
CC activity. The polynucleotides may also be used for recombinant expression
CC of the protein
XX
SQ Sequence 1168 BP; 276 A; 312 C; 300 G; 280 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.09e-92	Length:	1168
Score:	1045.00	Matches:	196
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	96.40%	Indels:	0
DB:	3	Gaps:	0

US-10-071-174-2 (1-204) x AAZ90049 (1-1168)

```
Qy          9 ThrThrMetAlaAspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGly 28
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGG 60

Qy         29 TyrCysAlaArgGluProGlyThrProGluProAlaProSerThrProGluAlaAlaVal 48
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACTGCGCCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTG 120

Qy         49 LeuArgSerAlaAlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeu 68
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTC 180

Qy         69 GlyTyrProGlyAsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAsp 88
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGAC 240

Qy         89 SerProGlyProThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeu 108
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTG 300

Qy        109 GluArgGlyProLeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLys 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAG 360

Qy        129 GluGlnGluGlyAspValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArg 148
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGG 420

Qy        149 LeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHis 168
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCAC 480

Qy        169 PhePheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeu 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTG 540

Qy        189 SerCysLeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeuLeu 204
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 588
```

RESULT 3
AAZ90039
ID AAZ90039 standard; cDNA; 582 BP.
XX
AC AAZ90039;
XX
DT 09-MAY-2000 (first entry)
XX
DE Hydrophobic domain containing protein clone HP02403 coding sequence.
XX
KW Hydrophobic domain; clone HP02403; nutritional supplement; SCID; HIV;
KW cell proliferation; immune stimulant; immune deficiency; tumour; pain;
KW rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
KW myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
KW anti-inflammatory; infection; bodily characteristic; ss.
XX
OS Homo sapiens.
XX
PN WO200000506-A2.
XX
PD 06-JAN-2000.
XX
PF 18-JUN-1999; 99WO-JP003242.
XX
PR 26-JUN-1998; 98JP-00180008.
XX
PA (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
PI Kato S, Kimura T;
XX
DR WPI; 2000-160665/14.
DR P-PSDB; AAY78802.
XX
PT Novel human proteins having hydrophobic domains used for research and
PT diagnostic purposes.
XX
PS Claim 3; Page 84; 117pp; English.
XX
CC This sequence represents the hydrophobic domain containing protein, clone
CC HP00631 coding region. The sequence is isolated from a human stomach
CC cancer cell line. The HP02403 protein contains one putative transmembrane
CC domain. The protein shows homology to the japanese quail apoptosis
CC regualtor NR-13. The invention relates to human proteins with hydrophobic
CC domains, the DNA and the cDNA encoding them. The polynucleotides and
CC proteins are predicted to have biological activities which make them
CC suitable for treating, preventing or ameliorating medical conditions in
CC humans and animals. Suggested activities include nutritional activity
CC (nutritional source or supplement); cytokine and cell
CC proliferation/differentiation activity; immune stimulating (e.g. as
CC vaccines) or suppressing activity (e.g. to treat various immune
CC deficiencies such as SCIDS or HIV, connective tissue disease, systemic
CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC and autoimmune inflammatory eye disease, as well as asthma, allergies and
CC organ transplantation); haematopoiesis regulating activity (e.g. in
CC treatment of myeloid or lymphoid cell deficiencies); tissue growth
CC activity (e.g. wound healing and tissue repair, ulcers, burns,
CC periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
CC activity; haemostatic and thrombolytic activity (e.g. treating
CC haemophilias); receptor/ligand activity; anti-inflammatory activity; and
CC tumour inhibition activity. The polynucleotides are also stated to be
CC useful for gene therapy. Other activities include inhibiting infections
CC caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
CC malaria); effecting bodily characteristics such as, e.g. weight, colour,
CC skin, effecting biorhythms or caricadic cycles; enhancing fertility;
CC treatment of depression; treatment of pain; hormonal or endocrine
CC activity. The polynucleotides may also be used for recombinant expression
CC of the protein
XX
SQ Sequence 582 BP; 85 A; 193 C; 191 G; 113 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.28e-92	Length:	582
Score:	1035.00	Matches:	194
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	95.48%	Indels:	0
DB:	3	Gaps:	0

```
US-10-071-174-2 (1-204) x AAZ90039 (1-582)

Qy          11 MetAlaAspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCys 30
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 60

Qy          31 AlaArgGluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArg 50
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 120

Qy          51 SerAlaAlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyr 70
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 180

Qy          71 ProGlyAsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerPro 90
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 240

Qy          91 GlyProThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArg 110
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 300

Qy         111 GlyProLeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGln 130
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 360

Qy         131 GluGlyAspValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeuMet 150
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 420

Qy         151 GlyGlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHisPhePhe 170
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 480

Qy         171 ArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeuSerCys 190
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 540

Qy         191 LeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeuLeu 204
               |||||||||||||||||||||||||||||||||||||||||||
Db         541 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 582

RESULT 4
AAH47022
ID   AAH47022 standard; cDNA; 522 BP.
XX
AC   AAH47022;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Human Bcl-2-like polypeptide encoding cDNA (clone HLIBE40).
XX
KW   Bcl-2-like polypeptide; autoimmune disorder; allergy; immunomodulatory;
KW   respiratory; cardiovascular; antiarthritic; immunostimulant; vaccine;
KW   immunosuppressive; antiinflammatory; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             44. .448
FT                   /*tag= a
FT                   /product= "Bcl-2-like polypeptide"
FT                   /note= "gene No. 2"
XX
PN   WO200157060-A1.
XX
PD   09-AUG-2001.
XX
PF   31-JAN-2001; 2001WO-US003080.
XX
PR   01-FEB-2000; 2000US-0179487P.
PR   07-FEB-2000; 2000US-0180697P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Duan DR,  Ni J;
XX
DR   WPI; 2001-476279/51.
DR   P-PSDB; AAB85666.
XX
PT   Nucleic acids encoding human Bcl-2-like polypeptides, useful for
PT   preventing, diagnosing and/or treating.
XX
PS   Claim 1; Page 276; 285pp; English.
XX
CC   The invention provides nucleic acid molecules (NAM1) encoding 4 human Bcl
CC   -2-like polypeptides (PEP1). The NAM1 and PEP1 may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate Bcl-2-like polypeptides' expression. The NAM1 may be used
CC   to produce the soluble Bcl-2-like polypeptides by standard recombinant
CC   methodolgy. The polypeptides may also be used as antigens in the
CC   production of antibodies against Bcl-2 and in assays to identify
CC   modulators of Bcl-2 expression and activity. The anti-Bcl-2 antibodies
CC   and antagonists may be used to down regulate expression and activity. The
CC   anti-PEP1 antibodies may also be used as diagnostic agents for detecting
CC   the presence of Bcl-2 polyPs in samples (e.g. by enzyme linked
CC   immunosorbant assay (ELISA)). Disorders that may be prevented, diagnosed
CC   and/or treated by the above methods include, immunodeficiencies (e.g. a
CC   gammaglobulinemia and B cell lymphoproliferative disorder), autoimmune
CC   disorders (e.g. rheumatoid arthritis and Grave's disease), allergic
CC   reactions, inflammations, respiratory diseases and cardiovascular
CC   disorders (a full list of disorders is given in the specification). The
CC   present sequence represents a human Bcl-2-like polypeptide encoding cDNA
XX
SQ   Sequence 522 BP; 102 A; 148 C; 154 G; 118 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.96e-59        Length:          522
Score:                    701.00          Matches:         130
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              64.67%          Indels:          0
DB:                       4               Gaps:            0

US-10-071-174-2 (1-204) x AAH47022 (1-522)

Qy          75 PheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerProGlyProThrTrp 94
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           3 TTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGG 62

Qy          95 GlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuVal 114
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 GGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTG 122

Qy         115 ThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspVal 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         123 ACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTC 182

Qy         135 AlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArg 154
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 GCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGC 242

Qy         155 AlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHisPhePheArgThrProPhe 174
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         243 GCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTT 302

Qy         175 ProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeuSerCysLeuLeuThrThr 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         303 CCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACA 362

Qy         195 AlaPheIleTyrLeuTrpThrArgLeuLeu 204
               ||||||||||||||||||||||||||||||
Db         363 GCCTTCATTTATCTCTGGACACGATTATTA 392

RESULT 5
ABK41913
```

ID ABK41913 standard; cDNA; 548 BP.
XX
AC ABK41913;
XX
DT 21-MAY-2002 (first entry)
XX
DE cDNA encoding novel human connective tissue related polypeptide #301.
XX
KW Human; connective tissue related disorder; cancer; gene therapy;
KW cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200155343-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001322.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-565190/63.
DR P-PSDB; AAU86735.
XX
PT Nucleic acid encoding novel connective tissue associated polypeptides,
PT used in diagnosing, preventing, treating or ameliorating a disorder such
PT as cancer or rheumatoid arthritis.
XX
PS Claim 4; SEQ ID NO 311; 673pp; English.
XX
CC The present invention relates to the isolation of novel human connective
CC tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
CC (cDNA and genomic) sequences encoding them. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of diseases associated with connective tissue(s), including
CC cancer. The polynucleotide sequences of the invention are also useful in
CC gene therapy. ABK41613-ABK42101 represent cDNA sequences encoding the
CC novel human connective tissue related polypeptides. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 548 BP; 99 A; 154 C; 167 G; 121 T; 0 U; 7 Other;

Alignment Scores:

Pred. No.:	2.47e-53	Length:	548
Score:	639.00	Matches:	127
Percent Similarity:	97.69%	Conservative:	0
Best Local Similarity:	97.69%	Mismatches:	3
Query Match:	58.95%	Indels:	2
DB:	4	Gaps:	0

US-10-071-174-2 (1-204) x ABK41913 (1-548)

```
Qy         75 PheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerProGlyProThrTrp 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 TTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGG 133

Qy         95 GlyArgValValThrLeuValThrPheAlaGlyThrLeuGluArgGlyProLeuVal 114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        134 GC-AGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTG 192

Qy        115 ThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspVal 134
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 ACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTC 252

Qy        135 AlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArg 154
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 GCCSGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGG 312

Qy        155 AlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHisPhePheArgThrProPhe 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 CC-TGGCTGCARGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTT 371

Qy        175 ProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeuSerCysLeuLeuThrThr 194
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 CCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACA 431

Qy        195 AlaPheIleTyrLeuTrpThrArgLeuLeu 204
              ||||||||||||||||||||||||||||||
```

Db 432 GCCTTCATTTATCTCTGGACACGATTATTA 461

RESULT 6
ADB59580
ID ADB59580 standard; cDNA; 548 BP.
XX
AC ADB59580;
XX
DT 04-DEC-2003 (first entry)
XX
DE Connective tissue related polynucleotide #301.
XX
KW cytostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;
KW antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;
KW antiinflammatory; antiallergic; antiasthmatic; dermatological;
KW nephrotopic; virucide; fungicide; antibacterial; antiparasitic;
KW gene therapy; ds; connective tissues disorder; rheumatoid arthritis;
KW systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;
KW cancer metastasis; neoplasia; leukaemia; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disease;
KW atherosclerosis; myocarditis; cardiopulmonary bypass complication;
KW autoimmune disease; multiple sclerosis; allergic reaction; asthma;
KW rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;
KW gastrointestinal disorder; inflammatory bowel disease;
KW organ transplant rejection; immune system disorder; Bruton's disease;
KW X-linked lymphoproliferative syndrome;
KW B-cell lymphoproliferative disorder; HIV; AIDS; infection;
KW chromosome identification; chromosome mapping;
KW connective tissue related polynucleotide; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003054375-A1.
XX
PD 20-MAR-2003.
XX
PF 07-MAR-2002; 2002US-00092154.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764847.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-634869/60.
DR P-PSDB; ADB60069.
XX
PT New connective tissue-related polypeptides and polynucleotides, useful
PT for treating, preventing and/or prognosing e.g. disorders of connective
PT tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or
PT neoplasias.
XX
PS Claim 1; SEQ ID NO 311; 248pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I), which
CC comprises a sequence that is at least 95 % identical to a connective
CC tissue-related polynucleotide encoding connective tissue antigens (CTA).
CC The polypeptide or polynucleotide is useful for preventing, treating, or
CC ameliorating medical conditions in a mammal. The connective tissue
CC polypeptides, polynucleotides and antibodies are particularly useful for
CC treating, preventing and/or prognosing disorders of connective tissues
CC (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,
CC scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or
CC neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g.
CC Alzheimer's disease, or Parkinson's disease), cardiovascular diseases
CC (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass
CC complications), autoimmune diseases (e.g. systemic lupus erythematosus,
CC rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.

Alignment Scores:

Pred. No.:	2.47e-53	Length:	548
Score:	639.00	Matches:	127
Percent Similarity:	97.69%	Conservative:	0
Best Local Similarity:	97.69%	Mismatches:	3
Query Match:	58.95%	Indels:	2
DB:	9	Gaps:	0

US-10-071-174-2 (1-204) x ADB59580 (1-548)

```
Qy         75 PheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerProGlyProThrTrp 94
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         74 TTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGG 133
```

```
Qy          95 GlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuVal 114
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         134 GC-AGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTG 192

Qy         115 ThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspVal 134
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         193 ACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTC 252

Qy         135 AlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArg 154
               |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         253 GCCSGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGG 312

Qy         155 AlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHisPhePheArgThrProPhe 174
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         313 CC-TGGCTGCARGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTT 371

Qy         175 ProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeuSerCysLeuLeuThrThr 194
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         372 CCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACA 431

Qy         195 AlaPheIleTyrLeuTrpThrArgLeuLeu 204
               ||||||||||||||||||||||||||||||
Db         432 GCCTTCATTTATCTCTGGACACGATTATTA 461
```

RESULT 7
ABQ44403
ID ABQ44403 standard; DNA; 874 BP.
XX
AC ABQ44403;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30994.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 874 BP; 374 A; 291 C; 90 G; 119 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.09e-45	Length:	874
Score:	559.00	Matches:	110
Percent Similarity:	83.75%	Conservative:	24
Best Local Similarity:	68.75%	Mismatches:	26
Query Match:	51.57%	Indels:	0
DB:	6	Gaps:	0

US-10-071-174-2 (1-204) x ABQ44403 (1-874)

```
Qy           1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
               :::::::::|||||||||:::|||||||||:::   ||||||||||||:::|||||||||
Db         149 ATAATTAACCAATTACGAAAACGCACCACCATAACCGACCCGCTACGAAAACGCACCGAA 208

Qy          21 LeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArgGluProGlyThrProGluProAla 40
               ||||||||   |||||||||  |||   ||||||:::|||   ||||||||||||
Db         209 CTATTACTAACCGACTACCTAAAATACTACGCCCGAAAACCCGACACCCCCGAACCGACG 268

Qy          41 ProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHis 60
               ||||||||||||||   ||||||||||||||||||   |||:::|||||||||||||||
Db         269 CCATCCACGCCCGAAACCGCCGTACTACGCTCCGCGACCGCCAAATTACGACAAATTCAC 328

Qy          61 ArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeu 80
               ||||||||||||||||||||||||   ||||||   ||||||||||||||||:::   |||
Db         329 CGATCCTTTTTCTCCGCCTACCTCGACTACCCCGAAAACCGCTTCGAACTAATAACGCTA 388

Qy          81 MetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLeu 100
               :::   :::|||||||||||||||:::|||   ||||||      :::::::::||||||
Db         389 ATAACGAATTCCGTACTCTCCGACAACCCCGACCCCACCTAAAACAAAATAATAACGCTC 448

Qy         101 ValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLys 120
               ||||||||||||   |||||||||::::::   ||||||:::|||||||||   ||||||
Db         449 GTAACCTTCGCAAAAACGCTACTAAAAAAAAAAACCGCTAATAACCGCCCGATAAAAAAAA 508

Qy         121 TrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArg 140
               ||||||||||||||||||:::|||:::   ||||||||||||:::   ||||||
Db         509 TAAAACTTCCAACCGCGACTAAAAAAAACAAAAAAAACGACGTCGCCCGAAACTACCAACGC 568

Qy         141 LeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGln 160
               |||:::   ||||||:::|||||||||:::   ||||||||||||   ||||||   |||
Db         569 CTAATAACCTTACTAAACTCGCGACTCATAAAACAACACCGCGCCTAACTACAAACTCAA 628
```

RESULT 8
ABQ44402/c
ID ABQ44402 standard; DNA; 874 BP.
XX
AC ABQ44402;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30993.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX

PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 874 BP; 119 A; 90 C; 291 G; 374 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.09e-45	Length:	874
Score:	559.00	Matches:	110
Percent Similarity:	83.75%	Conservative:	24
Best Local Similarity:	68.75%	Mismatches:	26
Query Match:	51.57%	Indels:	0
DB:	6	Gaps:	0

US-10-071-174-2 (1-204) x ABQ44402 (1-874)

```
Qy          1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
              ::::::::::||||||||||:::||||||||||:::   ||||||||||||:::||||||||
Db        726 ATAATTAACCAATTACGAAAACGCACCACCATAACCGACCCGCTACGAAAACGCACCGAA 667

Qy         21 LeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArgGluProGlyThrProGluProAla 40
              |||||||||   |||||||||   |||   ||||||:::|||   ||||||||||||
Db        666 CTATTACTAACCGACTACCTAAAATACTACGCCCGAAAACCCGACACCCCCGAACCGACG 607

Qy         41 ProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHis 60
              ||||||||||||||   ||||||||||||||||||   |||:::||||||||||||||||
Db        606 CCATCCACGCCCGAAACCGCCGTACTACGCTCCGCGACCGCCAAATTACGACAAATTCAC 547

Qy         61 ArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeu 80
              ||||||||||||||||||||||   ||||||   ||||||||||||||:::   |||
Db        546 CGATCCTTTTTCTCCGCCTACCTCGACTACCCCGAAAACCGCTTCGAACTAATAACGCTA 487

Qy         81 MetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLeu 100
              :::   :::||||||||||||||||:::|||   ||||||      :::::::::||||||
Db        486 ATAACGAATTCCGTACTCTCCGACAACCCCGACCCCACCTAAAACAAAATAATAACGCTC 427

Qy        101 ValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLys 120
              ||||||||||||   ||||||||::::::   ||||||:::|||||||||   ||||||   |||
Db        426 GTAACCTTCGCAAAAACGCTACTAAAAAAAAAACCGCTAATAACCGCCCGATAAAAAAAA 367

Qy        121 TrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArg 140
              ||||||||||||||||||:::|||:::   ||||||||||||:::   ||||||
Db        366 TAAAACTTCCAACCGCGACTAAAAAAACAAAAAAACGACGTCGCCCGAAACTACCAACGC 307

Qy        141 LeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGln 160
              |||:::   ||||||:::||||||||||:::   ||||||||||||   ||||||   |||
Db        306 CTAATAACCTTACTAAACTCGCGACTCATAAAACAACACCGCGCCTAACTACAAACTCAA 247
```

RESULT 9
ABQ44404
ID ABQ44404 standard; DNA; 874 BP.
XX
AC ABQ44404;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30995.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 874 BP; 144 A; 90 C; 320 G; 320 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9.92e-41	Length:	874
Score:	513.00	Matches:	111
Percent Similarity:	69.33%	Conservative:	2
Best Local Similarity:	68.10%	Mismatches:	50
Query Match:	47.32%	Indels:	0
DB:	6	Gaps:	0

US-10-071-174-2 (1-204) x ABQ44404 (1-874)

```
Qy          1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
              |||||||||   ||||||||||||     |||   |||   ||||||||||||   |||
Db        149 ATGGTTGATTAGTTGCGGGAGCGTATTATTATGGTCGATTCGTTGCGGGAGCGTATCGAG 208

Qy         21 LeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArgGluProGlyThrProGluProAla 40
              |||||||||   ||||||||||||||||||   ||||||   |||        |||   |||
Db        209 TTGTTGTTGGTCGATTATTTGGGGTATTGCGTTCGGGAATTCGGTATTTTCGAGTCGGCG 268

Qy         41 ProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHis 60
                 |||   |||       |||||||||   |||       |||||||||   |||:::
Db        269 TTATTTACGTTCGAGGTCGTCGTGTTGCGTTTCGCGGTCGTTAGGTTACGGTAGATTTAT 328

Qy         61 ArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeu 80
              |||   ||||||       |||   ||||||   ||||||||||||||||||||||||
Db        329 CGGTTTTTTTTTTTCGTTTATTTCGGTTATTTCGGGAATCGTTTCGAGTTGGTGGCGTTG 388

Qy         81 MetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLeu 100
              |||||||||   |||       ||||||   |||       ||||||||||||||||||
Db        389 ATGGCGGATTTCGTGTTTTTCGATAGTTTCGGTTTTATTTGGGGTAGAGTGGTGACGTTC 448

Qy        101 ValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLys 120
              |||   |||   ||||||||||||||||||||||   ||||||       ||||||||||||
Db        449 GTGATTTTCGTAGGGACGTTGTTGGAGAGAGGGTCGTTGGTGATCGTTCGGTGGAAGAAG 508

Qy        121 TrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArg 140
              |||||||||       ||||||||||||   ||||||||||||   |||||||||   |||
Db        509 TGGGGTTTTTAGTCGCGGTTAAAGGAGTAGGAGGGCGACGTCGTTCGGGATTGTTAGCGT 568

Qy        141 LeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGln 160
              ||||||   ||||||||||||||||   ||||||   :::|||   ||||||
Db        569 TTGGTGGTTTTGTTGAGTTCGCGGTTTATGGGGTAGTATCGCGTTTGGTTGTAGGTTTAG 628

Qy        161 GlyGlyTrp 163
              |||||||||
Db        629 GGCGGTTGG 637
```

RESULT 10
ABQ44405/c
ID ABQ44405 standard; DNA; 874 BP.
XX
AC ABQ44405;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30996.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 874 BP; 320 A; 320 C; 90 G; 144 T; 0 U; 0 Other;

Alignment Scores:			
Pred. No.:	9.92e-41	Length:	874
Score:	513.00	Matches:	111
Percent Similarity:	69.33%	Conservative:	2
Best Local Similarity:	68.10%	Mismatches:	50
Query Match:	47.32%	Indels:	0
DB:	6	Gaps:	0

US-10-071-174-2 (1-204) x ABQ44405 (1-874)

```
Qy          1 MetValAspGlnLeuArgGluArgThrThrMetAlaAspProLeuArgGluArgThrGlu 20
              |||||||||   ||||||||||||     |||   |||   ||||||||||||   |||
Db        726 ATGGTTGATTAGTTGCGGGAGCGTATTATTATGGTCGATTCGTTGCGGGAGCGTATCGAG 667

Qy         21 LeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArgGluProGlyThrProGluProAla 40
              |||||||||   ||||||||||||||||||   ||||||   |||        |||   |||
Db        666 TTGTTGTTGGTCGATTATTTGGGGTATTGCGTTCGGGAATTCGGTATTTTCGAGTCGGCG 607

Qy         41 ProSerThrProGluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHis 60
                 |||   |||       |||||||||   |||       |||||||||   |||:::
Db        606 TTATTTACGTTCGAGGTCGTCGTGTTGCGTTTCGCGGTCGTTAGGTTACGGTAGATTTAT 547

Qy         61 ArgSerPhePheSerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeu 80
              |||   ||||||       |||   ||||||   ||||||||||||||||||||||||
Db        546 CGGTTTTTTTTTTTCGTTTATTTCGGTTATTTCGGGAATCGTTTCGAGTTGGTGGCGTTG 487

Qy         81 MetAlaAspSerValLeuSerAspSerProGlyProThrTrpGlyArgValValThrLeu 100
              |||||||||   |||       ||||||   |||       ||||||||||||||||||
Db        486 ATGGCGGATTTCGTGTTTTTCGATAGTTTCGGTTTTATTTGGGGTAGAGTGGTGACGTTC 427

Qy        101 ValThrPheAlaGlyThrLeuLeuGluArgGlyProLeuValThrAlaArgTrpLysLys 120
              |||   |||   ||||||||||||||||||||||   ||||||       ||||||||||||
Db        426 GTGATTTTCGTAGGGACGTTGTTGGAGAGAGGGTCGTTGGTGATCGTTCGGTGGAAGAAG 367

Qy        121 TrpGlyPheGlnProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArg 140
              |||||||||       ||||||||||||   ||||||||||||   |||||||||   |||
Db        366 TGGGGTTTTTAGTCGCGGTTAAAGGAGTAGGAGGGCGACGTCGTTCGGGATTGTTAGCGT 307

Qy        141 LeuValAlaLeuLeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGln 160
              ||||||   ||||||||||||||||   ||||||   :::|||   ||||||
```

```
Db        306 TTGGTGGTTTTGTTGAGTTCGCGGTTTATGGGGTAGTATCGCGTTTGGTTGTAGGTTTAG 247

Qy        161 GlyGlyTrp 163
              |||||||||
Db        246 GGCGGTTGG 238

RESULT 11
ADD34132
ID   ADD34132 standard; DNA; 650 BP.
XX
AC   ADD34132;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Mouse mitochondrial DNA sequence SEQ ID NO:1910.
XX
KW   ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW   mitochondrial disease; oxidative phosphorylation dysfunction;
KW   oxidative stress; apoptosis; aging.
XX
OS   Mus musculus.
XX
PN   WO2003020220-A2.
XX
PD   13-MAR-2003.
XX
PF   30-AUG-2002; 2002WO-US027886.
XX
PR   30-AUG-2001; 2001US-0316323P.
PR   31-AUG-2001; 2001CA-02356540.
XX
PA   (UYEM-) UNIV EMORY.
XX
PI   Wallace DC,  Levy S,  Kerstann K,  Procaccio V;
XX
DR   WPI; 2003-300821/29.
XX
PT   Array containing probes for genes involved in mitochondrial biology,
PT   useful for determining mitochondrial biology gene expression profiles for
PT   use in diagnosing pathologies and identifying biochemical pathways.
XX
PS   Claim 2; SEQ ID NO 1910; 201pp; English.
XX
CC   The invention relates to a novel array comprising at least two isolated
CC   nucleotide molecules, each molecule having a sequence capable of uniquely
CC   hybridising to a nucleic acid molecule which is an expression product of
CC   a gene involved in mitochondrial biology. The array comprises two or more
CC   isolated nucleic acid molecules or spots, each molecule having a sequence
CC   chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC   of the invention is useful for determining an expression profile of a
CC   mouse or human sample containing nucleic acid, by contacting the array
CC   with the sample under conditions allowing selective hybridisation, and
CC   measuring hybridisation of nucleic acid in the sample to the array to
CC   produce an expression profile. The array is also useful for determining
CC   an expression profile of a first labelled sample containing nucleic acid
CC   relative to a second, differently labelled sample containing nucleic
CC   acid. The second sample is a reference or a standard. An array is useful
CC   for determining an expression profile diagnostic of an energy-metabolism-
CC   related physiological condition. An array of the invention is useful for
CC   determining mitochondrial biology gene expression profiles of organisms,
CC   such as human, mice and closely related species, tissue and organs of
CC   such organisms, which are useful for determining expression profiles
CC   diagnostic of energy metabolism-related physiological conditions,
CC   diagnosing such physiological conditions, identifying biochemical
CC   pathways, genes, and mutations involved in such physiological conditions,
CC   identifying therapeutic agents useful for preventing and/or treating such
CC   physiological conditions, evaluating and/or monitoring the efficacy of
CC   such therapies, and creating and identifying animal models of human
CC   energy metabolism-related physiological conditions. An array is also
CC   useful for defining expression signatures or profiles for mitochondrial
CC   diseases, as well as distinguishing clinical disorders that result from
CC   oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC   apoptosis and aging. An array of the invention contains probes of genes
CC   not previously recognised to participate in mitochondrial biology. The
CC   sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC   clones used to make the probes of the invention. Some sequences are not
CC   present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC   1906, 2408 and 2643.
XX
SQ   Sequence 650 BP; 143 A; 172 C; 170 G; 165 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.21e-37        Length:         650
Score:                  475.50          Matches:        94
Percent Similarity:     64.62%          Conservative:   32
Best Local Similarity:  48.21%          Mismatches:     56
Query Match:            43.87%          Indels:         13
DB:                     10              Gaps:           3

US-10-071-174-2 (1-204) x ADD34132 (1-650)

Qy         13 AspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArg 32
              |||||||||  |||||||||      ||||||:::||||||:::  :::||||||||||
Db         17 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 76

Qy         33 GluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArgSerAla 52
              ||||||  ||||||||||||  |||::::::  |||||||||:::||||||||||
Db         77 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGCCGAGGCGGCCTTGCTTCGCTCTGTG 136

Qy         53 AlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyrProGly 72
                  ::::::::|||   |||:::  |||||||||::::::             |||
Db        137 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 196

Qy         73 AsnArgPheGluLeuValAlaLeuMetAlaAspSerValLeuSerAspSerProGlyPro 92
              ||||||  |||||||||      |||||||||  :::||||||
Db        197 AATCGCCTGGAGCTGGTGAAACAGATGGCAGATAAGTTGCTCTCCAAAGACCAAGACTTC 256

Qy         93 ThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArgGlyPro 112
              :::|||  ::::::|||  |||:::  ||||||||||||||:::  :::||||||
Db        257 AGCTGGAGCCAACTGGTGATGCTCCTGGCCTTCGCGGGGACGCTTATGAATCAAGGCCCT 316

Qy        113 LeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGly 132
                 :::     :::  |||:::                                |||
Db        317 TACATGGCTGTCAAGCAGAAGAGG--------------------------GATCTGGGG 349

Qy        133 Asp---------ValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeu 149
              :::         |||  |||||||||      :::|||      |||  :::  |||
Db        350 AATCGTGTCATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTC 409

Qy        150 MetGly---GlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHis 168
              ||||||  :::||||||||  |||:::|||  ||||||||||||||||||||
Db        410 ATGGGGCGTCGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGGATGGCTTTTGCCGC 469

Qy        169 PhePheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeu 188
              ||||||:::  |||  ||||||  |||||||||:::  |||:::||||||||||||
Db        470 TTCTTCAAGAATCCTTTACCGCTCGGCTTCTGGAGAAGATTGCTGATTCAGGCTTTTCTG 529

Qy        189 SerCysLeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeu 203
              |||          ||||||      :::::|||  ||||||
Db        530 TCAGGCTTCTTTGCAACAGCCATCTTTTTTATCTGGAAACGTTTA 574

RESULT 12
ADD34582/c
ID   ADD34582 standard; DNA; 835 BP.
XX
AC   ADD34582;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Mouse mitochondrial DNA sequence SEQ ID NO:2360.
XX
KW   ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW   mitochondrial disease; oxidative phosphorylation dysfunction;
KW   oxidative stress; apoptosis; aging.
XX
OS   Mus musculus.
```

XX
PN WO2003020220-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027886.
XX
PR 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR WPI; 2003-300821/29.
XX
PT Array containing probes for genes involved in mitochondrial biology,
PT useful for determining mitochondrial biology gene expression profiles for
PT use in diagnosing pathologies and identifying biochemical pathways.
XX
PS Claim 2; SEQ ID NO 2360; 201pp; English.
XX
CC The invention relates to a novel array comprising at least two isolated
CC nucleotide molecules, each molecule having a sequence capable of uniquely
CC hybridising to a nucleic acid molecule which is an expression product of
CC a gene involved in mitochondrial biology. The array comprises two or more
CC isolated nucleic acid molecules or spots, each molecule having a sequence
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC of the invention is useful for determining an expression profile of a
CC mouse or human sample containing nucleic acid, by contacting the array
CC with the sample under conditions allowing selective hybridisation, and
CC measuring hybridisation of nucleic acid in the sample to the array to
CC produce an expression profile. The array is also useful for determining
CC an expression profile of a first labelled sample containing nucleic acid
CC relative to a second, differently labelled sample containing nucleic
CC acid. The second sample is a reference or a standard. An array is useful
CC for determining an expression profile diagnostic of an energy-metabolism-
CC related physiological condition. An array of the invention is useful for
CC determining mitochondrial biology gene expression profiles of organisms,
CC such as human, mice and closely related species, tissue and organs of
CC such organisms, which are useful for determining expression profiles
CC diagnostic of energy metabolism-related physiological conditions,
CC diagnosing such physiological conditions, identifying biochemical
CC pathways, genes, and mutations involved in such physiological conditions,
CC identifying therapeutic agents useful for preventing and/or treating such
CC physiological conditions, evaluating and/or monitoring the efficacy of
CC such therapies, and creating and identifying animal models of human
CC energy metabolism-related physiological conditions. An array is also
CC useful for defining expression signatures or profiles for mitochondrial
CC diseases, as well as distinguishing clinical disorders that result from
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC apoptosis and aging. An array of the invention contains probes of genes
CC not previously recognised to participate in mitochondrial biology. The
CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC clones used to make the probes of the invention. Some sequences are not
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC 1906, 2408 and 2643.
XX
SQ Sequence 835 BP; 238 A; 200 C; 185 G; 212 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.82e-13	Length:	835
Score:	232.50	Matches:	49
Percent Similarity:	56.52%	Conservative:	16
Best Local Similarity:	42.61%	Mismatches:	29
Query Match:	21.45%	Indels:	21
DB:	10	Gaps:	4

US-10-071-174-2 (1-204) x ADD34582 (1-835)

```
Qy         93 ThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGluArgGlyPro 112
              :::|||   |||                        |||||||||:::   :::|||||||
Db        826 TCCTGGCCTCGC----------------------GGGACACTTATGAATCAAGGCCCT 791

Qy        113 LeuValThrAlaArgTrpLysLysTrpGlyPheGlnProArgLeuLysGluGlnGluGly 132
                 :::      :::   |||:::                                |||
Db        790 TACATGGCTGTCAAGCAGAAGAGG-------------------------GATCTGGGG 758

Qy        133 Asp---------ValAlaArgAspCysGlnArgLeuValAlaLeuLeuSerSerArgLeu 149
                 :::         |||   |||||||||      :::|||      |||   :::   |||
Db        757 AATCGTGTCATAGTGACCCGAGACTGCTGTCTCATAGTGAACTTTCTGTATAATCTGCTC 698

Qy        150 MetGly---GlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAspGlyPheCysHis 168
              ||||||   :::|||||||||   |||:::|||   ||||||||||||||||||||
Db        697 ATGGGGCGTCGGCACCGCGCCAGGCTGGAGGCTCTCGGCGGCTGGGATGGCTTTTGCCGC 638

Qy        169 PhePheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuValGlnAlaPheLeu 188
              ||||||:::   |||   ||||||   |||||||||:::   |||:::|||||||||||||
Db        637 TTCTTCAAGAATCCTTTACCGCTCGGCTTCTGGAGAAGATTGCTGATTCAGGCTTTTCTG 578

Qy        189 SerCysLeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeu 203
              |||          ||||||      ::::::|||   ||||||
Db        577 TCAGGCTTCTTTGCAACAGCCATCTTTTTTATCTGGAAACGTTTA 533
```

RESULT 13
ADD34583
ID ADD34583 standard; DNA; 338 BP.
XX
AC ADD34583;
XX
DT 15-JAN-2004 (first entry)
XX
DE Mouse mitochondrial DNA sequence SEQ ID NO:2361.
XX
KW ds; mouse; array; mitochondrial; hybridisation; energy-metabolism;
KW mitochondrial disease; oxidative phosphorylation dysfunction;
KW oxidative stress; apoptosis; aging.
XX
OS Mus musculus.
XX
PN WO2003020220-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027886.
XX
PR 30-AUG-2001; 2001US-0316323P.
PR 31-AUG-2001; 2001CA-02356540.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Wallace DC, Levy S, Kerstann K, Procaccio V;
XX
DR WPI; 2003-300821/29.
XX
PT Array containing probes for genes involved in mitochondrial biology,
PT useful for determining mitochondrial biology gene expression profiles for
PT use in diagnosing pathologies and identifying biochemical pathways.
XX
PS Claim 2; SEQ ID NO 2361; 201pp; English.
XX
CC The invention relates to a novel array comprising at least two isolated
CC nucleotide molecules, each molecule having a sequence capable of uniquely
CC hybridising to a nucleic acid molecule which is an expression product of
CC a gene involved in mitochondrial biology. The array comprises two or more
CC isolated nucleic acid molecules or spots, each molecule having a sequence
CC chosen from sequence of 994 human probes and 2046 mouse probes. An array
CC of the invention is useful for determining an expression profile of a
CC mouse or human sample containing nucleic acid, by contacting the array
CC with the sample under conditions allowing selective hybridisation, and
CC measuring hybridisation of nucleic acid in the sample to the array to
CC produce an expression profile. The array is also useful for determining
CC an expression profile of a first labelled sample containing nucleic acid
CC relative to a second, differently labelled sample containing nucleic
CC acid. The second sample is a reference or a standard. An array is useful
CC for determining an expression profile diagnostic of an energy-metabolism-

CC related physiological condition. An array of the invention is useful for
CC determining mitochondrial biology gene expression profiles of organisms,
CC such as human, mice and closely related species, tissue and organs of
CC such organisms, which are useful for determining expression profiles
CC diagnostic of energy metabolism-related physiological conditions,
CC diagnosing such physiological conditions, identifying biochemical
CC pathways, genes, and mutations involved in such physiological conditions,
CC identifying therapeutic agents useful for preventing and/or treating such
CC physiological conditions, evaluating and/or monitoring the efficacy of
CC such therapies, and creating and identifying animal models of human
CC energy metabolism-related physiological conditions. An array is also
CC useful for defining expression signatures or profiles for mitochondrial
CC diseases, as well as distinguishing clinical disorders that result from
CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,
CC apoptosis and aging. An array of the invention contains probes of genes
CC not previously recognised to participate in mitochondrial biology. The
CC sequences shown in ADD33224-ADD35260 represent murine mitochondrial DNA
CC clones used to make the probes of the invention. Some sequences are not
CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,
CC 1906, 2408 and 2643.
XX
SQ Sequence 338 BP; 76 A; 95 C; 100 G; 67 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.7e-10	Length:	338
Score:	199.00	Matches:	38
Percent Similarity:	72.46%	Conservative:	12
Best Local Similarity:	55.07%	Mismatches:	19
Query Match:	18.36%	Indels:	0
DB:	10	Gaps:	0

US-10-071-174-2 (1-204) x ADD34583 (1-338)

```
Qy         13 AspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeuGlyTyrCysAlaArg 32
              |||||||||  |||||||||      ||||||:::||||||:::   :::||||||||||
Db        132 GACCCACTGCATGAACGCACTAGACGGCTGCTGTCTGACTACATATTCTTCTGCGCACGG 191

Qy         33 GluProGlyThrProGluProAlaProSerThrProGluAlaAlaValLeuArgSerAla 52
              ||||||  ||||||||||||   |||::::::  |||||||||:::||||||||||
Db        192 GAGCCGGACACCCCAGAGCCACCGCCCACGTCTGTCGAGGCGGCCTTGCTTCGCTCTGTG 251

Qy         53 AlaAlaArgLeuArgGlnIleHisArgSerPhePheSerAlaTyrLeuGlyTyrProGly 72
                 :::::::::|||   |||:::   |||||||||::::::           |||
Db        252 ACTAGGCAGATCCAGCAGGAGCACCAAGAATTTTTTTCCTCCTTCTGCGAAAGCCGGGGC 311

Qy         73 AsnArgPheGluLeuValAlaLeuMet 81
              ||||||  |||||||||     |||
Db        312 AATCGCCTGGAGCTGGTGAAACAGATG 338
```

RESULT 14
ABT09344
ID ABT09344 standard; DNA; 658 BP.
XX
AC ABT09344;
XX
DT 05-DEC-2002 (first entry)
XX
DE Phase-1 Rat CT gene SEQ ID No 432.
XX
KW Rat; toxicity study; rat toxic response gene; toxicological response;
KW drug development; phase-1 rat CT gene; ds.
XX
OS Rattus sp.
XX
PN WO200266682-A2.
XX
PD 29-AUG-2002.
XX
PF 29-JAN-2002; 2002WO-US002935.
XX
PR 29-JAN-2001; 2001US-0264933P.
PR 26-JUL-2001; 2001US-0308161P.
XX
PA (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY INC.
XX
PI Farris G, Hicken SH, Farr SB;
XX
DR WPI; 2002-674961/72.
XX
PT Evaluating the toxicity of an agent, useful in drug development or in
PT determining toxicological responses to a new drug, by determining the
PT expression of rat toxicologically relevant genes in the test animal in
PT response to the test agent.
XX
PS Disclosure; Page 213; 388pp; English.
XX
CC The invention relates to a method used for evaluating the toxicity of an
CC agent comprising determining the expression of a rat toxic response
CC gene(s) in the test animal in response to the agent. The method is useful
CC in drug development, particularly for conducting toxicity studies and
CC analysis before a new drug or compound is approved for human consumption
CC or use. The method is also useful in determining toxicological responses
CC to a new drug. This polynucleotide sequence represents a phase-1 rat CT
CC gene of the invention
XX
SQ Sequence 658 BP; 138 A; 178 C; 194 G; 147 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	0.00287	Length:	658
Score:	129.00	Matches:	58
Percent Similarity:	37.44%	Conservative:	24
Best Local Similarity:	26.48%	Mismatches:	75
Query Match:	11.90%	Indels:	62
DB:	6	Gaps:	12

US-10-071-174-2 (1-204) x ABT09344 (1-658)

```
Qy          8 ArgThrThrMetAlaAspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeu 27
              ||:::  :::     |||     |||            ||||||     ::::::
Db         51 CGCAGTGTGCTGAATTCGCCCTTCGCGGAATTCGGGGCCTTTTTGTTACAGGGTTTCATC 110

Qy         28 ---GlyTyrCysAlaArgGluProGly---ThrProGlu---ProAlaProSerThrPro 44
                     |||     |||   |||||||||                        |||
Db        111 CAGGATCGAGCAGAGAGGATGGCTGGGGAGACACCTGAGCTGACCTTGGAGCAGCCGCCC 170

Qy         45 GluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHisArgSerPhePhe 64
              :::  |||     :::     :::          ||||||:::|||
Db        171 CAGGACGCATCCACCAAGAAGCTGAGCGAGTGTCTCAGGCGAATTGGCGATGAACTG--- 227

Qy         65 SerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeuMetAlaAspSer 84
                                   |||     ||||||     ::::::||||||
Db        228 --------------------GACAACAACATGGAGCTGCAGAGGATGATTGCTGAT--- 263

Qy         85 ValLeuSerAspSerProGlyPro---------------------------------- 92
              |||  :::||||||||||
Db        264 GTGGATACAGACTCCCCCCGAGAGGTCTTCTTCCGTGTGGCAGCTGACATGTTTGCAGAC 323

Qy         93 ---------ThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGlu 109
                      |||||||||||||||   |||     ||||||     |||:::
Db        324 GGCAACTTCAACTGGGGCCGGGTGGTTGCCCTTTTCTACTTTGCTAGCAAACTGGTGCTC 383

Qy        110 ArgGly---------------ProLeuValThrAlaArgTrpLysLysTrpGlyPheGln 124
              :::                |||  :::                  |||
Db        384 AAGGCCCTGTGCACTAAAGTGCCCGAGCTGATCAGAACCATCATGGGCTGG--------- 434

Qy        125 ProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArgLeuValAlaLeu 144
                                                                  :::
Db        435 ------------------------------------------------ACACTGGACTTC 446

Qy        145 LeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAsp 164
              |||      ||||||:::             |||:::|||  ||||||||||||||||
Db        447 CTCCGGGAGCGGCTGCTT-----------GTCTGGATCCAAGACCAGGGTGGCTGGGAT 494

Qy        165 GlyPheCysHisPhePheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuVal 184
              |||         :::|||   ||||||              |||     |||   |||
Db        495 GGCCTCCTTTCCTACTTCGGGACCCCC-----------ACATGG------CAGACAGTG 536
```

```
Qy        185 GlnAlaPheLeuSerCysLeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeu 203
              |||::::::   :::|||   ||||||   :::   :::|||   :::|||
Db        537 ACCATCTTTGTGGCTGGAGTCCTC---ACTGCCTCGCTCACCATCTGGAAGAAGCTT 590

RESULT 15
ADG30848
ID   ADG30848 standard; DNA; 658 BP.
XX
AC   ADG30848;
XX
DT   26-FEB-2004  (first entry)
XX
DE   Liver toxicity predictive DNA 28.
XX
KW   liver toxicity; Norway rat; 24 hour combo All; ds.
XX
OS   Unidentified.
XX
PN   WO2003085083-A2.
XX
PD   16-OCT-2003.
XX
PF   01-APR-2003; 2003WO-US010141.
XX
PR   01-APR-2002; 2002US-0369287P.
XX
PA   (PHAS-) PHASE 1 MOLECULAR TOXICOLOGY.
XX
PI   Kier L,  Nolan TD,  Sankar U,  Derbel M;
XX
DR   WPI; 2003-804300/75.
XX
PT   Predicting the liver toxicity of an agent to an individual by using the
PT   test expression profile with a set of reference expression profiles in a
PT   Predictive Model to determine whether the agent will reduce liver
PT   toxicity in the individual.
XX
PS   Claim 2; Page 153; 379pp; English.
XX
CC   The invention relates to a novel method for predicting the liver toxicity
CC   of an agent to an individual comprising obtaining a biological sample
CC   from an individual treated with the agent, measuring the expression of
CC   one or more liver toxicity predictive genes in the sample and using the
CC   test expression profile with a set of reference expression profiles in a
CC   predictive model to determine whether the agent will reduce liver
CC   toxicity in the individual. The method of the invention may be useful for
CC   predicting the liver toxicity of an agent to an individual. The current
CC   sequence is that of the liver toxicity predictive DNA of the invention
CC   which represents a 24 hour combo All gene.
XX
SQ   Sequence 658 BP; 138 A; 178 C; 194 G; 147 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.:              0.00287       Length:         658
Score:                  129.00        Matches:        58
Percent Similarity:     37.44%        Conservative:   24
Best Local Similarity:  26.48%        Mismatches:     75
Query Match:            11.90%        Indels:         62
DB:                     10            Gaps:           12

US-10-071-174-2 (1-204) x ADG30848 (1-658)

Qy          8 ArgThrThrMetAlaAspProLeuArgGluArgThrGluLeuLeuLeuAlaAspTyrLeu 27
              |||:::   :::      |||   .|||           ||||||      ::::::
Db         51 CGCAGTGTGCTGAATTCGCCCTTCGCGGAATTCGGGGCCTTTTTGTTACAGGGTTTCATC 110

Qy         28 ---GlyTyrCysAlaArgGluProGly---ThrProGlu---ProAlaProSerThrPro 44
                  |||      |||   |||||||||                  |||
Db        111 CAGGATCGAGCAGAGAGGATGGCTGGGGAGACACCTGAGCTGACCTTGGAGCAGCCGCCC 170

Qy         45 GluAlaAlaValLeuArgSerAlaAlaAlaArgLeuArgGlnIleHisArgSerPhePhe 64
              :::   |||      :::      :::      ||||||:::|||

Db        171 CAGGACGCATCCACCAAGAAGCTGAGCGAGTGTCTCAGGCGAATTGGCGATGAACTG--- 227

Qy         65 SerAlaTyrLeuGlyTyrProGlyAsnArgPheGluLeuValAlaLeuMetAlaAspSer 84
                                        |||      ||||||      ::::::||||||
Db        228 --------------------GACAACAACATGGAGCTGCAGAGGATGATTGCTGAT--- 263

Qy         85 ValLeuSerAspSerProGlyPro---------------------------------- 92
              |||   :::||||||||
Db        264 GTGGATACAGACTCCCCCCGAGAGGTCTTCTTCCGTGTGGCAGCTGACATGTTTGCAGAC 323

Qy         93 ---------ThrTrpGlyArgValValThrLeuValThrPheAlaGlyThrLeuLeuGlu 109
                       ||||||||||||||||   |||      ||||||      |||:::
Db        324 GGCAACTTCAACTGGGGCCGGGTGGTTGCCCTTTTCTACTTTGCTAGCAAACTGGTGCTC 383

Qy        110 ArgGly---------------ProLeuValThrAlaArgTrpLysLysTrpGlyPheGln 124
              :::                 |||   :::                    |||
Db        384 AAGGCCCTGTGCACTAAAGTGCCCGAGCTGATCAGAACCATCATGGGCTGG--------- 434

Qy        125 ProArgLeuLysGluGlnGluGlyAspValAlaArgAspCysGlnArgLeuValAlaLeu 144
                                                                    :::
Db        435 ------------------------------------------------ACACTGGACTTC 446

Qy        145 LeuSerSerArgLeuMetGlyGlnHisArgAlaTrpLeuGlnAlaGlnGlyGlyTrpAsp 164
              |||      ||||||:::                ||:::|||   ||||||||||||||
Db        447 CTCCGGGAGCGGCTGCTT------------GTCTGGATCCAAGACCAGGGTGGCTGGGAT 494

Qy        165 GlyPheCysHisPheArgThrProPheProLeuAlaPheTrpArgLysGlnLeuVal 184
              |||         :::|||   ||||||              |||      |||   |||
Db        495 GGCCTCCTTTCCTACTTCGGGACCCCC------------ACATGG------CAGACAGTG 536

Qy        185 GlnAlaPheLeuSerCysLeuLeuThrThrAlaPheIleTyrLeuTrpThrArgLeu 203
              |||::::::   :::|||   ||||||   :::   :::|||   :::|||
Db        537 ACCATCTTTGTGGCTGGAGTCCTC---ACTGCCTCGCTCACCATCTGGAAGAAGCTT 590


Search completed: November  7, 2004, 06:26:07
Job time : 514 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 7, 2004, 01:20:07 ; Search time 6234 Seconds
(without alignments)
6728.579 Million cell updates/sec

Title: US-10-071-174-1
Perfect score: 887
Sequence: 1 cgggccaagaaaaccagcga.....ctcttccttgagtgaaagaa 887

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	887	100.0	887	9	AF326964	AF326964 Homo sapi
2	763	86.0	1168	6	BD233466	BD233466 Human pro
3	615	69.3	615	9	AF285092	AF285092 Homo sapi
4	585	66.0	585	9	HSA458330	AJ458330 Homo sapi
5	582	65.6	582	6	BD233456	BD233456 Human pro
c 6	539	60.8	93287	9	AC023906	AC023906 Homo sapi
7	439	49.5	726	6	CQ752105	CQ752105 Sequence
c 8	267	30.1	214669	2	AC018903	AC018903 Homo sapi
9	60	6.8	60	6	CQ545914	CQ545914 Sequence
10	47	5.3	214669	2	AC018903	AC018903 Homo sapi
c 11	22	2.5	5147	6	BD183336	BD183336 Novel gen
c 12	22	2.5	5147	6	BD171150	BD171150 Novel gen
c 13	22	2.5	5147	9	AB051441	AB051441 Homo sapi
14	22	2.5	21560	9	AY245248	AY245248 Homo sapi
15	22	2.5	60828	9	HS1191B2	AL022237 Human DNA
c 16	22	2.5	277861	2	HSAC000406	AC000406 Homo sapi
17	21	2.4	3011	10	BC030863	BC030863 Mus muscu
18	21	2.4	4353	10	AK173033	AK173033 Mus muscu
19	21	2.4	6190	10	BC043122	BC043122 Mus muscu
c 20	21	2.4	13635	1	AE005068	AE005068 Halobacte
c 21	21	2.4	95178	2	AL391556	AL391556 Homo sapi
c 22	21	2.4	120311	10	AC006945	AC006945 Mus muscu
23	21	2.4	155300	2	AL929459	AL929459 Danio rer
c 24	21	2.4	163574	2	AC138826	AC138826 Homo sapi
25	21	2.4	167273	10	AC083894	AC083894 Mus muscu
c 26	21	2.4	172144	2	AC136529	AC136529 Rattus no
c 27	21	2.4	172997	9	AC117516	AC117516 Homo sapi
c 28	21	2.4	190189	2	AC027181	AC027181 Homo sapi
c 29	21	2.4	224038	10	AL732625	AL732625 Mouse DNA
c 30	21	2.4	226457	2	AC127098	AC127098 Rattus no
c 31	21	2.4	231546	2	AC125302	AC125302 Rattus no
32	21	2.4	238290	2	AC112431	AC112431 Rattus no
33	21	2.4	288494	2	AC111749	AC111749 Rattus no
c 34	21	2.4	316892	2	AC096101	AC096101 Rattus no
c 35	20	2.3	284	11	G47512	G47512 Z25086_1 Ze
36	20	2.3	823	6	CQ780061	CQ780061 Sequence
37	20	2.3	823	6	CQ782081	CQ782081 Sequence
38	20	2.3	823	6	BD124770	BD124770 Primer fo
39	20	2.3	823	6	BD126790	BD126790 Primer fo
40	20	2.3	1134	8	AK066814	AK066814 Oryza sat
41	20	2.3	1209	10	AF102501	AF102501 Mus muscu
42	20	2.3	1225	10	AF067660	AF067660 Mus muscu
43	20	2.3	1257	10	BC052690	BC052690 Mus muscu
c 44	20	2.3	1672	9	AY167728	AY167728 Pongo pyg
45	20	2.3	38674	9	HSU81831	U81831 Human cosmi

ALIGNMENTS

RESULT 1
AF326964
LOCUS AF326964 887 bp mRNA linear PRI 01-MAY-2001
DEFINITION Homo sapiens BCLB (BCLB) mRNA, complete cds.
ACCESSION AF326964
VERSION AF326964.1 GI:13898393
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS Ke,N., Godzik,A. and Reed,J.C.
TITLE Bcl-B, a novel Bcl-2 family member that differentially binds and regulates Bax and Bak
JOURNAL J. Biol. Chem. 276 (16), 12481-12484 (2001)
MEDLINE 21201065
PUBMED 11278245
REFERENCE 2 (bases 1 to 887)
AUTHORS Ke,N., Godzik,A. and Reed,J.C.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2000) The Burnham Institute, 10901 N. Torrey Pines Rd., La Jolla, CA 92037, USA
FEATURES Location/Qualifiers
source 1. .887
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q21"
gene 1. .887
/gene="BCLB"
CDS 50. .664
/gene="BCLB"
/note="BCL2 family protein; Boo-like"
/codon_start=1
/product="BCLB"
/protein_id="AAK48715.1"
/db_xref="GI:13898394"
/translation="MVDQLRERTTMADPLRERTELLLADYLGYCAREPGTPEPAPSTP
EAAVLRSAAARLRQIHRSFFSAYLGYPGNRFELVALMADSVLSDSPGPTWGRVVTLVT
FAGTLLERGPLVTARWKKWGFQPRLKEQEGDVARDCQRLVALLSSRLMGQHRAWLQAQ

```
                    GGWDGFCHFFRTPFPLAFWRKQLVQAFLSCLLTTAFIYLWTRLL"
ORIGIN

  Query Match           100.0%;  Score 887;  DB 9;  Length 887;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  887;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 60

Qy         61 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 120

Qy        121 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 180

Qy        181 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 240

Qy        241 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 300

Qy        301 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 360

Qy        361 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 420

Qy        421 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 480

Qy        481 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGA 540

Qy        541 TGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGT 600

Qy        601 CCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATT 660

Qy        661 ATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAG 720

Qy        721 ATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGA 780

Qy        781 ATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACT 840

Qy        841 CAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887


RESULT 2
BD233466
LOCUS       BD233466                1168 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION   BD233466
VERSION     BD233466.1  GI:33043236
KEYWORDS    JP 2002519016-A/12.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1168)
  AUTHORS   Kato,S. and Kimura,T.
  TITLE     Human protein having hydrophobic domain and DNA encoding the same
  JOURNAL   Patent: JP 2002519016-A 12 02-JUL-2002;
            SAGAMI CHEMICAL RESEARCH CENTER,PROTEGENE INC
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002519016-A/12
            PD   02-JUL-2002
            PF   18-JUN-1999 JP 2000557267
            PI   SEISHI KATO,TOMOKO KIMURA
            PC
            C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
             00
            CC   Human protein having hydrophobic domain and DNA encoding the
            CC        same
            FH   Key             Location/Qualifiers
            FT   source          1. .1168
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .1168
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           86.0%;  Score 763;  DB 6;  Length 1168;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  813;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         74 ACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGG 133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGG 60

Qy        134 TACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTG 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTG 120

Qy        194 CTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTC 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTC 180

Qy        254 GGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGAC 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGAC 240

Qy        314 AGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTG 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTG 300

Qy        374 GAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAG 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAG 360

Qy        434 GAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGG 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGG 420

Qy        494 CTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCAC 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCAC 480

Qy        554 TTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTG 613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTG 540

Qy        614 TCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAA 673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        541 TCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAA 600

Qy        674 CTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACA 733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACA 660

Qy        734 AGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGA 793
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        661 AGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTACCTGAATGCATACAAGGA 720

Qy        794 GTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGAC 853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGAC 780

Qy        854 AAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
              |||||||||||||||||||||||||||||||||||
Db        781 AAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 814
```

```
RESULT 3
AF285092
LOCUS       AF285092                615 bp    mRNA    linear   PRI 08-NOV-2001
DEFINITION  Homo sapiens Bcl-2-like protein 10 mRNA, complete cds.
ACCESSION   AF285092
VERSION     AF285092.1  GI:9837265
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 615)
  AUTHORS   Zhang,H., Holzgreve,W. and De Geyter,C.
  TITLE     Bcl2-L-10, a novel anti-apoptotic member of the Bcl-2 family,
            blocks apoptosis in the mitochondria death pathway but not in the
            death receptor pathway
  JOURNAL   Hum. Mol. Genet. 10 (21), 2329-2339 (2001)
  MEDLINE   21548034
   PUBMED   11689480
REFERENCE   2  (bases 1 to 615)
  AUTHORS   Zhang,H.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUL-2000) University Women's Hospital, Schanzenstra
            46, Basel 4057, Switzerland
FEATURES             Location/Qualifiers
     source          1. .615
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21"
                     /tissue_type="ovary"
     CDS             1. .615
                     /codon_start=1
                     /product="Bcl-2-like protein 10"
                     /protein_id="AAG00503.1"
                     /db_xref="GI:9837266"
                     /translation="MVDQLRERTTMADPLRERTELLLADYLGYCAREPGTPEPAPSTP
                     EAAVLRSAAARLRQIHRSFFSAYLGYPGNRFELVALMADSVLSDSPGPTWGRVVTLVT
                     FAGTLLERGPLVTARWKKWGFQPRLKEQEGDVARDCQRLVALLSSRLMGQHRAWLQAQ
                     GGWDGFCHFFRTPFPLAFWRKQLVQAFLSCLLTTAFIYLWTRLL"
ORIGIN

  Query Match           69.3%;  Score 615;  DB 9;  Length 615;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  615;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         50 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 60

Qy        110 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 120

Qy        170 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 180

Qy        230 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 240

Qy        290 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 300

Qy        350 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 360

Qy        410 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 420

Qy        470 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 480

Qy        530 GGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGA 589
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGA 540

Qy        590 AAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGG 649
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGG 600

Qy        650 ACACGATTATTATGA 664
              |||||||||||||||
Db        601 ACACGATTATTATGA 615
```

```
RESULT 4
HSA458330
LOCUS       HSA458330               585 bp    DNA     linear   PRI 25-APR-2002
DEFINITION  Homo sapiens NRH gene for anti-apoptotic protein.
ACCESSION   AJ458330
VERSION     AJ458330.1  GI:20338765
KEYWORDS    anti-apoptotic protein; NRH gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Aouacheria,A., Arnaud,E., Venet,S., Lalle,P., Gouy,M., Rigal,D. and
            Gillet,G.
  TITLE     Nrh, a human homologue of Nr-13 associates with Bcl-Xs and is an
            inhibitor of apoptosis
  JOURNAL   Oncogene 20 (41), 5846-5855 (2001)
  MEDLINE   21477277
   PUBMED   11593390
REFERENCE   2  (bases 1 to 585)
  AUTHORS   Gillet,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-APR-2002) Gillet G., Ibcp, CNRS UMR 5086, 7 PASSAGE
            DU VERCORS, 69367, FRANCE
FEATURES             Location/Qualifiers
     source          1. .585
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
     gene            1. .585
                     /gene="NRH"
     CDS             1. .585
                     /gene="NRH"
                     /function="apoptosis inhibitor"
```

```
                     /codon_start=1
                     /product="anti-apoptotic protein"
                     /protein_id="CAD30221.1"
                     /db_xref="GI:20338766"
                     /db_xref="GOA:Q9HD36"
                     /db_xref="Swiss-Prot:Q9HD36"
                     /translation="MADPLRERTELLLADYLGYCAREPGTPEPAPSTPEAAVLRSAAA
                     RLRQIHRSFFSAYLGYPGNRFELVALMADSVLSDSPGPTWGRVVTLVTFAGTLLERGP
                     LVTARWKKWGFQPRLKEQEGDVARDCQRLVALLSSRLMGQHRAWLQAQGGWDGFCHFF
                     RTPFPLAFWRKQLVQAFLSCLLTTAFIYLWTRLL"
ORIGIN

  Query Match             66.0%;  Score 585;  DB 9;  Length 585;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  585;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         80 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 60

Qy        140 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 199
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 120

Qy        200 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 259
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 180

Qy        260 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 319
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 240

Qy        320 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 379
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 300

Qy        380 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 439
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 360

Qy        440 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 499
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 420

Qy        500 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 559
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 480

Qy        560 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 619
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 540

Qy        620 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGA 664
               |||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGA 585


RESULT 5
BD233456
LOCUS       BD233456                582 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Human protein having hydrophobic domain and DNA encoding the same.
ACCESSION   BD233456
VERSION     BD233456.1  GI:33043226
KEYWORDS    JP 2002519016-A/2.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 582)
  AUTHORS   Kato,S. and Kimura,T.
  TITLE     Human protein having hydrophobic domain and DNA encoding the same
  JOURNAL   Patent: JP 2002519016-A 2 02-JUL-2002;
            SAGAMI CHEMICAL RESEARCH CENTER,PROTEGENE INC
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002519016-A/2
            PD   02-JUL-2002
            PF   18-JUN-1999 JP 2000557267
            PI   SEISHI KATO,TOMOKO KIMURA
            PC
            C12N15/09,C07K14/47,C12N1/15,C12N1/19,C12N5/10,C12N15/00,C12N5/ PC
             00
            CC   Human protein having hydrophobic domain and DNA encoding the
            CC        same
            FH   Key             Location/Qualifiers
            FT   source          1. .582
            FT                   /organism='Homo sapiens (human)'.
FEATURES             Location/Qualifiers
     source          1. .582
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             65.6%;  Score 582;  DB 6;  Length 582;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  582;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         80 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 60

Qy        140 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 199
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 120

Qy        200 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 259
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 180

Qy        260 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 319
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 240

Qy        320 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 379
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 300

Qy        380 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 439
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 360

Qy        440 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 499
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 420

Qy        500 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 559
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 480

Qy        560 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 619
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 540

Qy        620 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 661
               |||||||||||||||||||||||||||||||||||||||||
Db        541 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 582


RESULT 6
AC023906/c
LOCUS       AC023906                93287 bp    DNA     linear   PRI 06-JUL-2001
DEFINITION  Homo sapiens chromosome 15 clone CTD-2184D3 map 15q21.2, complete
            sequence.
ACCESSION   AC023906
VERSION     AC023906.7  GI:14595770
KEYWORDS    HTG.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-2000) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
REFERENCE   4  (bases 1 to 93287)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
COMMENT     On Jul 4, 2001 this sequence version replaced gi:12248292.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399

            Note:  Data from overlapping BACs AC010674 [drafting center:
            UWMSC], AC090970 [drafting center: UWMSC], and AC016824 [drafting
            center: GTC] was added for finishing
            ------------.
FEATURES             Location/Qualifiers
     source          1. .93287
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21.2"
                     /clone="CTD-2184D3"
                     /clone_lib="Cal Tech Human BAC library D"
                     /note="Data from overlapping clones CTD-2650P22 AC090970,
                     RP11-47K1 AC016824, and RP11-430B1 was added and the
                     consensus was determined from CTD-2184D3 to the extent
                     possible"
     misc_feature    1. .8893
                     /note="overlap with CTD-2650P22 AC090970"
     misc_feature    1. .8893
                     /note="overlap with RP11-47K1 AC016824"
     unsure          6078. .6140
                     /note="low quality data"
     unsure          10529. .10532
                     /note="low quality data"
     unsure          14722. .14808
                     /note="low quality data"
     unsure          37627
                     /note="low quality data"
     unsure          71552. .71593
                     /note="low quality data"
     misc_feature    82012. .93287
                     /note="overlap with RP11-430B1 AC010674"
ORIGIN
```

```
  Query Match             60.8%;  Score 539;  DB 9;  Length 93287;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-297;
  Matches  539;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64991 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 64932

Qy         61 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64931 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 64872

Qy        121 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64871 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 64812

Qy        181 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64811 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 64752

Qy        241 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64751 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 64692

Qy        301 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64691 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 64632

Qy        361 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64631 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 64572

Qy        421 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64571 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 64512

Qy        481 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGG 539
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64511 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGG 64453
```

```
RESULT 7
CQ752105
LOCUS       CQ752105                726 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 38039 from Patent WO02068579.
ACCESSION   CQ752105
VERSION     CQ752105.1  GI:42387450
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 38039 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             49.5%;  Score 439;  DB 6;  Length 726;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-239;
  Matches  489;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         50 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGTTGACCAGTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAG 60

Qy        110 CTGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 169
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGGTTGCTGGCCGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCG 120

Qy        170 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCATCCACGCCCGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCAC 180

Qy        230 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGTCCTTTTTCTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTG 240

Qy        290 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTC 300

Qy        350 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAG 360

Qy        410 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGC 420

Qy        470 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAG 480

Qy        530 GGCGGCTGGG 539
              ||||||||||
Db        481 GGCGGCTGGG 490


RESULT 8
AC018903/c
LOCUS       AC018903               214669 bp    DNA     linear   HTG 04-JUN-2000
DEFINITION  Homo sapiens chromosome 15 clone RP11-337B11 map 15q21, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC018903
VERSION     AC018903.2  GI:8247797
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 214669)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., Kaur,A.,
            Madan,A., Nesbitt,R., Shaffer,T. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 214669)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-DEC-1999) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
COMMENT     On Jun 4, 2000 this sequence version replaced gi:6630517.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399
            ------------.
            * NOTE: This record contains 192 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *        1      806: contig of 806 bp in length
            *      807      906: gap of unknown length
            *      907     2068: contig of 1162 bp in length
            *     2069     2168: gap of unknown length
            *     2169     3020: contig of 852 bp in length
            *     3021     3120: gap of unknown length
            *     3121     4284: contig of 1164 bp in length
            *     4285     4384: gap of unknown length
            *     4385     5201: contig of 817 bp in length
            *     5202     5301: gap of unknown length
            *     5302     6544: contig of 1243 bp in length
            *     6545     6644: gap of unknown length
            *     6645     7472: contig of 828 bp in length
            *     7473     7572: gap of unknown length
            *     7573     8756: contig of 1184 bp in length
            *     8757     8856: gap of unknown length
            *     8857     9658: contig of 802 bp in length
            *     9659     9758: gap of unknown length
            *     9759    10905: contig of 1147 bp in length
            *    10906    11005: gap of unknown length
            *    11006    11811: contig of 806 bp in length
            *    11812    11911: gap of unknown length
            *    11912    13293: contig of 1382 bp in length
            *    13294    13393: gap of unknown length
            *    13394    14211: contig of 818 bp in length
            *    14212    14311: gap of unknown length
            *    14312    15465: contig of 1154 bp in length
            *    15466    15565: gap of unknown length
            *    15566    16379: contig of 814 bp in length
            *    16380    16479: gap of unknown length
            *    16480    17624: contig of 1145 bp in length
            *    17625    17724: gap of unknown length
            *    17725    18578: contig of 854 bp in length
            *    18579    18678: gap of unknown length
            *    18679    19774: contig of 1096 bp in length
            *    19775    19874: gap of unknown length
            *    19875    20700: contig of 826 bp in length
            *    20701    20800: gap of unknown length
            *    20801    22036: contig of 1236 bp in length
            *    22037    22136: gap of unknown length
            *    22137    23023: contig of 887 bp in length
            *    23024    23123: gap of unknown length
            *    23124    24196: contig of 1073 bp in length
            *    24197    24296: gap of unknown length
            *    24297    25141: contig of 845 bp in length
            *    25142    25241: gap of unknown length
            *    25242    26385: contig of 1144 bp in length
            *    26386    26485: gap of unknown length
            *    26486    27325: contig of 840 bp in length
            *    27326    27425: gap of unknown length
            *    27426    28691: contig of 1266 bp in length
            *    28692    28791: gap of unknown length
            *    28792    29655: contig of 864 bp in length
            *    29656    29755: gap of unknown length
            *    29756    30881: contig of 1126 bp in length
```

```
*    30882    30981: gap of unknown length
*    30982    31813: contig of 832 bp in length
*    31814    31913: gap of unknown length
*    31914    33056: contig of 1143 bp in length
*    33057    33156: gap of unknown length
*    33157    34007: contig of 851 bp in length
*    34008    34107: gap of unknown length
*    34108    35320: contig of 1213 bp in length
*    35321    35420: gap of unknown length
*    35421    36207: contig of 787 bp in length
*    36208    36307: gap of unknown length
*    36308    37521: contig of 1214 bp in length
*    37522    37621: gap of unknown length
*    37622    38420: contig of 799 bp in length
*    38421    38520: gap of unknown length
*    38521    39726: contig of 1206 bp in length
*    39727    39826: gap of unknown length
*    39827    40635: contig of 809 bp in length
*    40636    40735: gap of unknown length
*    40736    41845: contig of 1110 bp in length
*    41846    41945: gap of unknown length
*    41946    42730: contig of 785 bp in length
*    42731    42830: gap of unknown length
*    42831    43917: contig of 1087 bp in length
*    43918    44017: gap of unknown length
*    44018    44825: contig of 808 bp in length
*    44826    44925: gap of unknown length
*    44926    46114: contig of 1189 bp in length
*    46115    46214: gap of unknown length
*    46215    47031: contig of 817 bp in length
*    47032    47131: gap of unknown length
*    47132    48292: contig of 1161 bp in length
*    48293    48392: gap of unknown length
*    48393    49176: contig of 784 bp in length
*    49177    49276: gap of unknown length
*    49277    50389: contig of 1113 bp in length
*    50390    50489: gap of unknown length
*    50490    51333: contig of 844 bp in length
*    51334    51433: gap of unknown length
*    51434    52765: contig of 1332 bp in length
*    52766    52865: gap of unknown length
*    52866    53710: contig of 845 bp in length
*    53711    53810: gap of unknown length
*    53811    54965: contig of 1155 bp in length
*    54966    55065: gap of unknown length
*    55066    55910: contig of 845 bp in length
*    55911    56010: gap of unknown length
*    56011    57558: contig of 1548 bp in length
*    57559    57658: gap of unknown length
*    57659    58509: contig of 851 bp in length
*    58510    58609: gap of unknown length
*    58610    59745: contig of 1136 bp in length
*    59746    59845: gap of unknown length
*    59846    60655: contig of 810 bp in length
*    60656    60755: gap of unknown length
*    60756    61976: contig of 1221 bp in length
*    61977    62076: gap of unknown length
*    62077    62896: contig of 820 bp in length
*    62897    62996: gap of unknown length
*    62997    64128: contig of 1132 bp in length
*    64129    64228: gap of unknown length
*    64229    65026: contig of 798 bp in length
*    65027    65126: gap of unknown length
*    65127    66269: contig of 1143 bp in length
*    66270    66369: gap of unknown length
*    66370    67220: contig of 851 bp in length
*    67221    67320: gap of unknown length
*    67321    68425: contig of 1105 bp in length
*    68426    68525: gap of unknown length
*    68526    69322: contig of 797 bp in length
*    69323    69422: gap of unknown length
*    69423    70568: contig of 1146 bp in length
*    70569    70668: gap of unknown length
*    70669    71547: contig of 879 bp in length
*    71548    71647: gap of unknown length
*    71648    73629: contig of 1982 bp in length
*    73630    73729: gap of unknown length
*    73730    74587: contig of 858 bp in length
*    74588    74687: gap of unknown length
*    74688    75780: contig of 1093 bp in length
*    75781    75880: gap of unknown length
*    75881    77196: contig of 1316 bp in length
*    77197    77296: gap of unknown length
*    77297    78651: contig of 1355 bp in length
*    78652    78751: gap of unknown length
*    78752    79557: contig of 806 bp in length
*    79558    79657: gap of unknown length
*    79658    80774: contig of 1117 bp in length
*    80775    80874: gap of unknown length
*    80875    81690: contig of 816 bp in length
*    81691    81790: gap of unknown length
*    81791    82928: contig of 1138 bp in length
*    82929    83028: gap of unknown length
*    83029    83902: contig of 874 bp in length
*    83903    84002: gap of unknown length
*    84003    85193: contig of 1191 bp in length
*    85194    85293: gap of unknown length
*    85294    86149: contig of 856 bp in length

  Query Match           30.1%;  Score 267;  DB 2;  Length 214669;
  Best Local Similarity  100.0%;  Pred. No. 8.3e-141;
  Matches  267;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        621 TGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTAA 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168497 TGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTAA 168438

Qy        681 CCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACTG 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168437 CCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACTG 168378

Qy        741 AGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGAGTCCTGA 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168377 AGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGAGTCCTGA 168318

Qy        801 GGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGACAAGAATG 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168317 GGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGACAAGAATG 168258

Qy        861 CAAATGGCTCTTCCTTGAGTGAAAGAA 887
              ||||||||||||||||||||||||||
Db     168257 CAAATGGCTCTTCCTTGAGTGAAAGAA 168231


RESULT 9
CQ545914
LOCUS       CQ545914                 60 bp    DNA     linear   PAT 30-JAN-2004
DEFINITION  Sequence 15549 from Patent WO0210449.
ACCESSION   CQ545914
VERSION     CQ545914.1  GI:41512178
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Faigler,S.
  TITLE     Oligonucleotide library for detecting rna transcripts and splice
            variants that populate a transcriptome
  JOURNAL   Patent: WO 0210449-A 15549 07-FEB-2002;
            Compugen Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
```

ORIGIN

Query Match 6.8%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.2e-22;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        565 CCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTT 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTT 60
```

RESULT 10
AC018903
LOCUS AC018903 214669 bp DNA linear HTG 04-JUN-2000
DEFINITION Homo sapiens chromosome 15 clone RP11-337B11 map 15q21, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC018903
VERSION AC018903.2 GI:8247797
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 214669)
AUTHORS Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B., Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., Kaur,A., Madan,A., Nesbitt,R., Shaffer,T. and Hood,L.
TITLE Sequencing of human chromosome 15 D15S146-D15S117 region
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 214669)
AUTHORS Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B., Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G., James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T. and Hood,L.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-1999) Multimegabase Sequencing Center, University of Washington, PO BOX 357730, Seattle, WA 98195, USA
COMMENT On Jun 4, 2000 this sequence version replaced gi:6630517.
---------- Genome Center
Center: Multimegabase Sequencing Center
Center code: UWMSC
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leerowen@systemsbiology.org
---------- Summary Statistics
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399
----------.
* NOTE: This record contains 192 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

*	1	806: contig of 806 bp in length
*	807	906: gap of unknown length
*	907	2068: contig of 1162 bp in length
*	2069	2168: gap of unknown length
*	2169	3020: contig of 852 bp in length
*	3021	3120: gap of unknown length
*	3121	4284: contig of 1164 bp in length
*	4285	4384: gap of unknown length
*	4385	5201: contig of 817 bp in length
*	5202	5301: gap of unknown length
*	5302	6544: contig of 1243 bp in length
*	6545	6644: gap of unknown length
*	6645	7472: contig of 828 bp in length
*	7473	7572: gap of unknown length
*	7573	8756: contig of 1184 bp in length
*	8757	8856: gap of unknown length
*	8857	9658: contig of 802 bp in length
*	9659	9758: gap of unknown length
*	9759	10905: contig of 1147 bp in length
*	10906	11005: gap of unknown length
*	11006	11811: contig of 806 bp in length
*	11812	11911: gap of unknown length
*	11912	13293: contig of 1382 bp in length
*	13294	13393: gap of unknown length
*	13394	14211: contig of 818 bp in length
*	14212	14311: gap of unknown length
*	14312	15465: contig of 1154 bp in length
*	15466	15565: gap of unknown length
*	15566	16379: contig of 814 bp in length
*	16380	16479: gap of unknown length
*	16480	17624: contig of 1145 bp in length
*	17625	17724: gap of unknown length
*	17725	18578: contig of 854 bp in length
*	18579	18678: gap of unknown length
*	18679	19774: contig of 1096 bp in length
*	19775	19874: gap of unknown length
*	19875	20700: contig of 826 bp in length
*	20701	20800: gap of unknown length
*	20801	22036: contig of 1236 bp in length
*	22037	22136: gap of unknown length
*	22137	23023: contig of 887 bp in length
*	23024	23123: gap of unknown length
*	23124	24196: contig of 1073 bp in length
*	24197	24296: gap of unknown length
*	24297	25141: contig of 845 bp in length
*	25142	25241: gap of unknown length
*	25242	26385: contig of 1144 bp in length
*	26386	26485: gap of unknown length
*	26486	27325: contig of 840 bp in length
*	27326	27425: gap of unknown length
*	27426	28691: contig of 1266 bp in length
*	28692	28791: gap of unknown length
*	28792	29655: contig of 864 bp in length
*	29656	29755: gap of unknown length
*	29756	30881: contig of 1126 bp in length
*	30882	30981: gap of unknown length
*	30982	31813: contig of 832 bp in length
*	31814	31913: gap of unknown length
*	31914	33056: contig of 1143 bp in length
*	33057	33156: gap of unknown length
*	33157	34007: contig of 851 bp in length
*	34008	34107: gap of unknown length
*	34108	35320: contig of 1213 bp in length
*	35321	35420: gap of unknown length
*	35421	36207: contig of 787 bp in length
*	36208	36307: gap of unknown length
*	36308	37521: contig of 1214 bp in length
*	37522	37621: gap of unknown length
*	37622	38420: contig of 799 bp in length
*	38421	38520: gap of unknown length
*	38521	39726: contig of 1206 bp in length
*	39727	39826: gap of unknown length
*	39827	40635: contig of 809 bp in length
*	40636	40735: gap of unknown length
*	40736	41845: contig of 1110 bp in length
*	41846	41945: gap of unknown length
*	41946	42730: contig of 785 bp in length
*	42731	42830: gap of unknown length
*	42831	43917: contig of 1087 bp in length
*	43918	44017: gap of unknown length
*	44018	44825: contig of 808 bp in length
*	44826	44925: gap of unknown length
*	44926	46114: contig of 1189 bp in length
*	46115	46214: gap of unknown length
*	46215	47031: contig of 817 bp in length
*	47032	47131: gap of unknown length

```
*    47132    48292: contig of 1161 bp in length
*    48293    48392: gap of unknown length
*    48393    49176: contig of 784 bp in length
*    49177    49276: gap of unknown length
*    49277    50389: contig of 1113 bp in length
*    50390    50489: gap of unknown length
*    50490    51333: contig of 844 bp in length
*    51334    51433: gap of unknown length
*    51434    52765: contig of 1332 bp in length
*    52766    52865: gap of unknown length
*    52866    53710: contig of 845 bp in length
*    53711    53810: gap of unknown length
*    53811    54965: contig of 1155 bp in length
*    54966    55065: gap of unknown length
*    55066    55910: contig of 845 bp in length
*    55911    56010: gap of unknown length
*    56011    57558: contig of 1548 bp in length
*    57559    57658: gap of unknown length
*    57659    58509: contig of 851 bp in length
*    58510    58609: gap of unknown length
*    58610    59745: contig of 1136 bp in length
*    59746    59845: gap of unknown length
*    59846    60655: contig of 810 bp in length
*    60656    60755: gap of unknown length
*    60756    61976: contig of 1221 bp in length
*    61977    62076: gap of unknown length
*    62077    62896: contig of 820 bp in length
*    62897    62996: gap of unknown length
*    62997    64128: contig of 1132 bp in length
*    64129    64228: gap of unknown length
*    64229    65026: contig of 798 bp in length
*    65027    65126: gap of unknown length
*    65127    66269: contig of 1143 bp in length
*    66270    66369: gap of unknown length
*    66370    67220: contig of 851 bp in length
*    67221    67320: gap of unknown length
*    67321    68425: contig of 1105 bp in length
*    68426    68525: gap of unknown length
*    68526    69322: contig of 797 bp in length
*    69323    69422: gap of unknown length
*    69423    70568: contig of 1146 bp in length
*    70569    70668: gap of unknown length
*    70669    71547: contig of 879 bp in length
*    71548    71647: gap of unknown length
*    71648    73629: contig of 1982 bp in length
*    73630    73729: gap of unknown length
*    73730    74587: contig of 858 bp in length
*    74588    74687: gap of unknown length
*    74688    75780: contig of 1093 bp in length
*    75781    75880: gap of unknown length
*    75881    77196: contig of 1316 bp in length
*    77197    77296: gap of unknown length
*    77297    78651: contig of 1355 bp in length
*    78652    78751: gap of unknown length
*    78752    79557: contig of 806 bp in length
*    79558    79657: gap of unknown length
*    79658    80774: contig of 1117 bp in length
*    80775    80874: gap of unknown length
*    80875    81690: contig of 816 bp in length
*    81691    81790: gap of unknown length
*    81791    82928: contig of 1138 bp in length
*    82929    83028: gap of unknown length
*    83029    83902: contig of 874 bp in length
*    83903    84002: gap of unknown length
*    84003    85193: contig of 1191 bp in length
*    85194    85293: gap of unknown length
*    85294    86149: contig of 856 bp in length

  Query Match             5.3%;  Score 47;  DB 2;  Length 214669;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         67 GGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGT 113
              |||||||||||||||||||||||||||||||||||||||||||||||
Db      74826 GGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGT 74872


RESULT 11
BD183336/c
LOCUS       BD183336                5147 bp    DNA     linear   PAT 17-JUN-2003
DEFINITION  Novel genes and proteins encoded by the genes.
ACCESSION   BD183336
VERSION     BD183336.1  GI:31875536
KEYWORDS    JP 2002345492-A/49.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5147)
  AUTHORS   Ohara,O., Nagase,T. and Nakajima,D.
  TITLE     Novel genes and proteins encoded by the genes
  JOURNAL   Patent: JP 2002345492-A 49 03-DEC-2002;
            KAZUSA DNA RESEARCH INSTITUTE
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002345492-A/49
            PD   03-DEC-2002
            PF   26-FEB-2002 JP 2002049009
            PI   OSAMU OHARA,TAKAHIRO NAGASE,DAISUKE NAKAJIMA
            PC   C12N15/09,C07K14/47//A61K31/711,A61K38/00,A61K48/00,A61P25/00,
            PC   A61P25/14,
            PC   A61P25/18,A61P35/00,C12N15/00,A61K37/02
            CC   Novel genes and proteins encoded by the genes FH   Key
                Location/Qualifiers
            FT   CDS              (2074). .(2346).
FEATURES            Location/Qualifiers
     source         1. .5147
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
ORIGIN

  Query Match             2.5%;  Score 22;  DB 6;  Length 5147;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        512 GCCTGGCTGCAGGCTCAGGGCG 533
              ||||||||||||||||||||||
Db       4330 GCCTGGCTGCAGGCTCAGGGCG 4309


RESULT 12
BD171150/c
LOCUS       BD171150                5147 bp    DNA     linear   PAT 17-JAN-2003
DEFINITION  Novel gene and protein encoded thereby.
ACCESSION   BD171150
VERSION     BD171150.1  GI:27876962
KEYWORDS    WO 02052005-A/6.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5147)
  AUTHORS   Ohara,O., Nagase,T. and Nakajima,D.
  TITLE     Novel gene and protein encoded thereby
  JOURNAL   Patent: WO 02052005-A 6 04-JUL-2002;
            KAZUSA DNA RESEARCH INSTITUTE,OSAMU OHARA,TAKAHIRO NAGASE, DAISUKE
            NAKAJIMA
COMMENT     OS   Homo sapiens (human)
            PN   WO 02052005-A/6
            PD   04-JUL-2002
            PF   20-DEC-2001 WO 2001JP011217
            PR   22-DEC-2000 JP 00P   389742
            PI   OSAMU OHARA,TAKAHIRO NAGASE,DAISUKE NAKAJIMA
            PC   C12N15/12,C07K14/47
            CC   Novel gene and protein encoded thereby
            FH   Key              Location/Qualifiers
```

```
           FT   CDS             (2074). .(2346).
FEATURES             Location/Qualifiers
     source          1. .5147
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             2.5%;  Score 22;  DB 6;  Length 5147;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        512 GCCTGGCTGCAGGCTCAGGGCG 533
              ||||||||||||||||||||||
Db       4330 GCCTGGCTGCAGGCTCAGGGCG 4309


RESULT 13
AB051441/c
LOCUS       AB051441                5147 bp    mRNA    linear   PRI 06-OCT-2001
DEFINITION  Homo sapiens mRNA for KIAA1654 protein, partial cds.
ACCESSION   AB051441
VERSION     AB051441.1  GI:13359180
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Hirosawa,M., Nagase,T., Murahashi,Y., Kikuno,R. and Ohara,O.
  TITLE     Identification of novel transcribed sequences on human chromosome
            22 by expressed sequence tag mapping
  JOURNAL   DNA Res. 8 (1), 1-9 (2001)
  MEDLINE   21156230
   PUBMED   11258795
REFERENCE   2  (bases 1 to 5147)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-2000) Osamu Ohara, Kazusa DNA Research Institute,
            Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
            Fax:81-438-52-3914)
FEATURES             Location/Qualifiers
     source          1. .5147
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fh10218"
                     /note="vector:pBluescriptII SK plus"
     gene            1. .5147
                     /gene="KIAA1654"
     CDS             <2074. .2349
                     /gene="KIAA1654"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /product="KIAA1654 protein"
                     /protein_id="BAB33324.1"
                     /db_xref="GI:13359181"
                     /translation="AGLRAQICMGKVWSQGVEGGEPAQEGPCSSDLLLWSLLATPPAQ
                     GRPSPVSSYLAPALNALLVSSLQWSSLQPNLKVAASTLTLSLLCSAP"
ORIGIN

  Query Match             2.5%;  Score 22;  DB 9;  Length 5147;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        512 GCCTGGCTGCAGGCTCAGGGCG 533
              ||||||||||||||||||||||
Db       4330 GCCTGGCTGCAGGCTCAGGGCG 4309


RESULT 14
AY245248
LOCUS       AY245248                21560 bp    DNA     linear   PRI 12-MAR-2003
DEFINITION  Homo sapiens BCL2-interacting killer (apoptosis-inducing) (BIK)
            gene, complete cds.
ACCESSION   AY245248
VERSION     AY245248.1  GI:28932927
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 21560)
  AUTHORS   Rieder,M.J., Livingston,R.J., Daniels,M.R., Chung,M.-W.,
            Miyamoto,K.E., Nguyen,C.P., Nguyen,D.A., Poel,C.L., Robertson,P.D.,
            Schackwitz,W.S., Sherwood,J.K., Witrak,L.A. and Nickerson,D.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2003) Genome Sciences, University of Washington,
            1705 NE Pacific, Seattle, WA 98195, USA
COMMENT     To cite this work please use:  NIEHS-SNPs, Environmental Genome
            Project, NIEHS ES15478, Department of Genome Sciences, Seattle, WA
            (URL: http://egp.gs.washington.edu).
FEATURES             Location/Qualifiers
     source          1. .21560
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     repeat_region   102. .155
                     /rpt_family="L2"
                     /rpt_type=dispersed
     repeat_region   210. .270
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     variation       221
                     /frequency="0.05"
                     /replace="t"
     variation       222
                     /frequency="0.02"
                     /replace="a"
     variation       307
                     /frequency="0.19"
                     /replace="g"
     variation       353
                     /frequency="0.01"
                     /replace="a"
     repeat_region   408. .718
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     variation       594
                     /frequency="0.20"
                     /replace="a"
     variation       620
                     /frequency="0.01"
                     /replace="t"
     repeat_region   764. .879
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     variation       845
                     /frequency="0.24"
                     /replace="g"
     variation       858
                     /frequency="0.01"
                     /replace="a"
     variation       1006
                     /frequency="0.01"
                     /replace="a"
     variation       1011
                     /frequency="0.01"
                     /replace="t"
     variation       1270
                     /frequency="0.01"
                     /replace="c"
     variation       1313
                     /frequency="0.01"
```

```
                     /replace="t"
     variation       1415
                     /frequency="0.45"
                     /replace="t"
     repeat_region   1488. .1596
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     variation       1785
                     /frequency="0.17"
                     /replace="t"
     variation       1828
                     /frequency="0.01"
                     /replace="a"
     variation       1903
                     /frequency="0.27"
                     /replace="a"
     gene            1907. .20871
                     /gene="BIK"
     mRNA            join(1907. .1959,15172. .15339,18854. .18952,19655. .19784,
                     20372. .20871)
                     /gene="BIK"
                     /product="BCL2-interacting killer (apoptosis-inducing)"
     variation       2234
                     /gene="BIK"
                     /frequency="0.01"
                     /replace="t"
     variation       2388
                     /gene="BIK"
                     /frequency="0.01"
                     /replace="a"
     variation       2424
                     /gene="BIK"
                     /frequency="0.01"
                     /replace="c"
     variation       2954
                     /gene="BIK"
                     /frequency="0.01"
                     /replace="g"
     repeat_region   3088. .3211
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     variation       3095
                     /gene="BIK"
                     /frequency="0.01"
                     /replace="t"
     variation       3168
                     /gene="BIK"
                     /frequency="0.01"
                     /replace="t"
     variation       3186. .3189
                     /gene="BIK"
                     /frequency="0.01"
                     /replace=""
     misc_feature    3268. .9489
                     /gene="BIK"
                     /note="Region not scanned for variation"
     repeat_region   3346. .3654
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   3660. .3743
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   3781. .4068
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   4078. .4377
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   4466. .4524
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   4581. .4717
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   5477. .5756
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   6618. .6918
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   6921. .7212
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   7445. .7568
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   7578. .7620
                     /rpt_family="MER2_type"
                     /rpt_type=dispersed
     repeat_region   7621. .7912
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   7913. .8097
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   8515. .8825
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   8830. .9128
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   9149. .9263
                     /rpt_family="Alu"
                     /rpt_type=dispersed
     repeat_region   9589. .9700
                     /rpt_family="MIR"
                     /rpt_type=dispersed
     variation       9874
                     /gene="BIK"
                     /frequency="0.02"
                     /replace="a"
     variation       9897
                     /gene="BIK"
                     /frequency="0.01"
                     /replace="t"
     variation       10011
                     /gene="BIK"
                     /frequency="0.06"
                     /replace="g"
     variation       10030
                     /gene="BIK"
                     /frequency="0.02"
                     /replace="a"
     variation       10036
                     /gene="BIK"
                     /frequency="0.09"
                     /replace="t"
     variation       10160

  Query Match             2.5%;  Score 22;  DB 9;  Length 21560;
  Best Local Similarity   100.0%;  Pred. No. 2.7;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        512 GCCTGGCTGCAGGCTCAGGGCG 533
              ||||||||||||||||||||||
Db       9659 GCCTGGCTGCAGGCTCAGGGCG 9680


RESULT 15
HS1191B2
LOCUS       HS1191B2               60828 bp    DNA     linear   PRI 05-JUN-2003
DEFINITION  Human DNA sequence from clone CTB-1191B2 on chromosome
            22q13.2-13.3, complete sequence.
ACCESSION   AL022237
VERSION     AL022237.1  GI:3845439
KEYWORDS    HTG.
```

```
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 60828)
  AUTHORS   Cobley,V.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 5, 1998 this sequence version replaced gi:3820993.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep CTB-1191B2 is
            from the human BAC library described in U-J. Kim et al. (1996)
            Genomics 34, 213-218.
            VECTOR: pBeloBAC11
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 22, constructed by the Sanger Centre Chromosome 22
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr22
            IMPORTANT: This sequence is not the entire insert of clone
            CTB-1191B2 It may be shorter because we sequence overlapping
            sections only once, except for a short overlap.
            The true left end of clone CTB-1191B2 is at 1 in this sequence. The
            true right end of clone RP3-323M22 is at 18859 in this sequence.
FEATURES             Location/Qualifiers
     source          1. .60828
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="q13.2-13.3"
                     /clone="CTB-1191B2"
                     /clone_lib="CIT978SK-B"
     misc_feature    1. .340
                     /note="match: GSS: Em:B13896"
     misc_feature    7. .225
                     /note="match: GSS: Em:B14210"
     repeat_region   complement(142. .453)
                     /note="AluY repeat: matches 2. .304 of consensus"
     repeat_region   complement(456. .563)
                     /note="MIR repeat: matches 158. .256 of consensus"
     repeat_region   564. .889
                     /note="THE1B repeat: matches 1. .364 of consensus"
     repeat_region   complement(890. .936)
                     /note="MIR repeat: matches 110. .158 of consensus"
     repeat_region   1063. .1100
                     /note="2.4 copies 16 mer GGGCTGGCTCAGGAGC 67% conserved"
     repeat_region   1129. .1139
                     /note="2.8 copies 4 mer CAGC 22% conserved"
     repeat_region   1264. .1273
                     /note="2.5 copies 4 mer TAAA 20% conserved"
     repeat_region   1722. .1732
                     /note="2.8 copies 4 mer ACAG 22% conserved"
     repeat_region   1793. .1802
                     /note="2.5 copies 4 mer GCCT 20% conserved"
     repeat_region   1846. .1855
                     /note="2.5 copies 4 mer GGAA 20% conserved"
     repeat_region   1864. .2371
                     /note="L2 repeat: matches 2493. .3134 of consensus"
     repeat_region   2398. .2455
                     /note="MER5A1 repeat: matches 1. .68 of consensus"
     repeat_region   2458. .2723
                     /note="AluSx repeat: matches 37. .303 of consensus"
     repeat_region   2735. .3034
                     /note="AluSx repeat: matches 1. .300 of consensus"
     repeat_region   3045. .3346
                     /note="AluSx repeat: matches 1. .306 of consensus"
     repeat_region   complement(3504. .3821)
                     /note="AluY repeat: matches 1. .311 of consensus"
     misc_feature    4231. .4718
                     /note="CpG island"
                     /evidence=not_experimental
     repeat_region   4585. .4599
                     /note="3.0 copies 5 mer CCCGC 30% conserved"
     repeat_region   4935. .5010
                     /note="AluYd3a1 repeat: matches 224. .299 of consensus"
     repeat_region   5044. .5135
                     /note="AluY repeat: matches 2. .93 of consensus"
     repeat_region   5136. .5257
                     /note="AluSg/x repeat: matches 177. .298 of consensus"
     repeat_region   complement(5337. .5422)
                     /note="MIR repeat: matches 13. .102 of consensus"
     repeat_region   5427. .5436
                     /note="3.3 copies 3 mer GGT 20% conserved"
     repeat_region   5755. .6070
                     /note="AluJo repeat: matches 13. .312 of consensus"
     repeat_region   6081. .6378
                     /note="AluSx repeat: matches 1. .298 of consensus"
     repeat_region   complement(6380. .6477)
                     /note="MER45 repeat: matches 2. .101 of consensus"
     repeat_region   6577. .6874
                     /note="AluSc repeat: matches 1. .298 of consensus"
     repeat_region   complement(6875. .7116)
                     /note="L1MB2 repeat: matches 5913. .6159 of consensus"
     repeat_region   7119. .7251
                     /note="FLAM_C repeat: matches 1. .133 of consensus"
     repeat_region   7252. .7267
                     /note="5.3 copies 3 mer AGA 32% conserved"
     repeat_region   7263. .7274
                     /note="3.0 copies 4 mer AAGA 24% conserved"
     repeat_region   7275. .7287
                     /note="3.2 copies 4 mer GAAG 26% conserved"
     repeat_region   7285. .7301
                     /note="4.2 copies 4 mer AGGG 34% conserved"
     gene            7348. .7712
                     /gene="bK1191B2.1"
                     /pseudo
     CDS             7348. .7712
                     /gene="bK1191B2.1"
                     /note="bK1191B2.1 (40S Ribososmal Protein S25 pseudogene)
                     match: proteins: Wp:CE04691 Sw:P25111 Sw:P46301 Sw:Q03409
                     Tr:O15597 Sw:P48588 Tr:O74172 Sw:P07282 Sw:P79009"
                     /pseudo
                     /codon_start=1
                     /evidence=experimental
     repeat_region   7770. .8080
                     /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   complement(8110. .8299)
                     /note="AluJo repeat: matches 121. .309 of consensus"
     repeat_region   complement(8300. .8428)
                     /note="L1MB2 repeat: matches 5767. .5893 of consensus"
     repeat_region   8429. .8589
```

```
                     /note="L1PA5 repeat: matches 5994. .6154 of consensus"
     repeat_region   8595. .8890
                     /note="AluSx repeat: matches 1. .296 of consensus"
     repeat_region   complement(8891. .8973)
                     /note="L1MB2 repeat: matches 5684. .5764 of consensus"
     repeat_region   complement(8974. .9110)
                     /note="AluJo repeat: matches 2. .133 of consensus"
     repeat_region   complement(9111. .9328)
                     /note="L1MB2 repeat: matches 5450. .5684 of consensus"
     repeat_region   complement(9330. .9400)
                     /note="MIR repeat: matches 80. .150 of consensus"
     repeat_region   9598. .9607
                     /note="2.5 copies 4 mer GCCA 20% conserved"
     repeat_region   9733. .9746
                     /note="3.5 copies 4 mer CCCT 21% conserved"
     repeat_region   10003. .10315
                     /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   10441. .10451
                     /note="2.2 copies 5 mer CCCTT 22% conserved"
     repeat_region   10640. .10765
                     /note="MIR repeat: matches 101. .233 of consensus"
     repeat_region   10832. .10847
                     /note="2.0 copies 8 mer AGAGTGTG 32% conserved"
     repeat_region   10843. .11068
                     /note="113.0 copies 2 mer GT 141% conserved"
     repeat_region   10853. .10899
                     /note="2.2 copies 21 mer GTGTGAGTACATGGGTGTTGT 60%
                     conserved"
     repeat_region   10887. .11068
                     /note="16.5 copies 11 mer GGTGTGTGTGT 114% conserved"
     repeat_region   11135. .11150
                     /note="8.0 copies 2 mer TA 32% conserved"
     repeat_region   11215. .11521
                     /note="AluYe5 repeat: matches 1. .309 of consensus"
     repeat_region   11523. .11886
                     /note="LTR10C repeat: matches 215. .562 of consensus"
     repeat_region   11899. .11917
                     /note="9.5 copies 2 mer TG 29% conserved"
     repeat_region   12600. .12893
                     /note="AluSp repeat: matches 1. .303 of consensus"
     repeat_region   12894. .13201
                     /note="AluYc2 repeat: matches 1. .309 of consensus"
     repeat_region   13205. .13220
                     /note="5.3 copies 3 mer AAG 32% conserved"
     repeat_region   13224. .13241
                     /note="2.2 copies 8 mer CTGAACCT 27% conserved"
     repeat_region   13547. .13688
                     /note="FRAM repeat: matches 23. .164 of consensus"

  Query Match             2.5%;  Score 22;  DB 9;  Length 60828;
  Best Local Similarity   100.0%;  Pred. No. 2.8;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        512 GCCTGGCTGCAGGCTCAGGGCG 533
              ||||||||||||||||||||||
Db      33682 GCCTGGCTGCAGGCTCAGGGCG 33703


Search completed: November  7, 2004, 04:47:20
Job time : 6238 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 6, 2004, 23:36:01 ; Search time 686 Seconds
(without alignments)
6787.517 Million cell updates/sec

Title: US-10-071-174-1
Perfect score: 887
Sequence: 1 cgggccaagaaaaccagcga.....ctcttccttgagtgaaagaa 887

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	887	100.0	887	6	AAD46683	Aad46683 Human Bcl
	2	763	86.0	1168	3	AAZ90049	Aaz90049 Hydrophob
	3	582	65.6	582	3	AAZ90039	Aaz90039 Hydrophob
	4	488	55.0	522	4	AAH47022	Aah47022 Human Bcl
	5	225	25.4	548	4	ABK41913	Abk41913 cDNA enco
	6	225	25.4	548	9	ADB59580	Adb59580 Connectiv
	7	60	6.8	60	6	ABN42801	Abn42801 Human spl
c	8	24	2.7	24	6	AAD46685	Aad46685 Human Bcl
	9	23	2.6	23	6	AAD46684	Aad46684 Human Bcl
	10	22	2.5	22	6	AAD46689	Aad46689 Human Bcl
	11	22	2.5	28	6	AAD46686	Aad46686 Human Bcl
c	12	22	2.5	34	6	AAD46688	Aad46688 Human Bcl
	13	22	2.5	874	6	ABQ44403	Abq44403 Oligonucl
c	14	22	2.5	874	6	ABQ44402	Abq44402 Oligonucl
c	15	22	2.5	5147	6	ABN83959	Abn83959 Human gen
c	16	21	2.4	30	6	AAD46687	Aad46687 Human Bcl
	17	21	2.4	3827	12	ADM66989	Adm66989 Murine ad
	18	20	2.3	650	10	ADD34132	Add34132 Mouse mit
	19	20	2.3	823	4	AAK91741	Aak91741 Human cDN
	20	20	2.3	823	4	AAK93761	Aak93761 Human cDN
	21	20	2.3	823	12	ADL28168	Adl28168 5' end of
	22	20	2.3	823	12	ADL30188	Adl30188 3' end of
c	23	20	2.3	835	10	ADD34582	Add34582 Mouse mit
c	24	20	2.3	90798	12	ADP68859	Adp68859 Human chr
	25	19	2.1	65	6	ABN30228	Abn30228 Rat splic
	26	19	2.1	387	3	ADF56850	Adf56850 Urogenita
c	27	19	2.1	507	11	ABD15534	Abd15534 Pseudomon
	28	19	2.1	665	6	ABT09478	Abt09478 Phase-1 R
	29	19	2.1	665	10	ADG30917	Adg30917 Liver tox
	30	19	2.1	665	12	ADG45503	Adg45503 Liver inf
c	31	19	2.1	936	11	ABD15622	Abd15622 Pseudomon
	32	19	2.1	1630	2	AAQ28270	Aaq28270 Encodes r
	33	19	2.1	1630	10	ABT41845	Abt41845 Toxicity
	34	19	2.1	1779	5	AAS78125	Aas78125 DNA encod
	35	19	2.1	3419	4	AAA91196	Aaa91196 Mouse IGF
	36	19	2.1	5561	6	ABI99223	Abi99223 Mouse isc
c	37	19	2.1	9588	10	ADG37082	Adg37082 Vector pP
c	38	19	2.1	28198	10	ADG37080	Adg37080 Mouse pla
	39	19	2.1	33053	6	ABQ67005	Abq67005 Human ang
	40	19	2.1	143391	10	ADL13648	Adl13648 Osteoarth
c	41	18	2.0	421	8	ABX54335	Abx54335 Bovine ES
c	42	18	2.0	474	9	ACH18221	Ach18221 Human adu
	43	18	2.0	530	3	AAF15017	Aaf15017 Trichoder
	44	18	2.0	539	8	ABX91658	Abx91658 Murine ge
	45	18	2.0	553	4	AAH13606	Aah13606 Human cDN

ALIGNMENTS

RESULT 1
AAD46683
ID AAD46683 standard; DNA; 887 BP.
XX
AC AAD46683;
XX
DT 27-JAN-2003 (first entry)
XX
DE Human Bcl-B DNA.
XX
KW Human; Bcl-2; Bcl-B; therapy; apoptosis; cell degenerative disorder;
KW proliferative disorder; muscle degeneration; Alzheimer's disease; CJD;
KW Creutzfeldt-Jacob's disease; Machado-Joseph disease; MJD; transgenic;
KW Parkinson's disease; Huntington's disease; HD; spinocerebellar ataxia;
KW SCA; dentatorubropallidoluysian atrophy; DRPLA; Kennedy's disease;
KW stroke; ischaemia; head trauma; neoplasia; anticonvulsant; vulnerary;
KW nootropic; neuroprotective; cytostatic; immunosuppressive; vasotropic;
KW cerebroprotective; autoimmune disorder; chromosome 15; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 50. .664
FT /*tag= a
FT /product= "Human Bcl-B protein"
XX
PN WO200272601-A2.
XX
PD 19-SEP-2002.
XX
PF 07-FEB-2002; 2002WO-US003547.
XX
PR 07-FEB-2001; 2001US-0267166P.
PR 07-FEB-2002; 2002US-00071174.
XX
PA (BURN-) BURNHAM INST.
XX
PI Reed JC, Ke N, Godzik A;
XX
DR WPI; 2002-723312/78.
DR P-PSDB; AAE29097.
XX
PT New isolated or recombinant Bcl-B nucleic acids and polypeptides, for
PT treating a disorder associated with apoptosis, such as cell degenerative

```
PT   or proliferative disorder e.g. cancer, Alzheimer's disease or Parkinson's
PT   disease.
XX
PS   Claim 11; Page 74; 82pp; English.
XX
CC   The invention relates to human member of Bcl-2 family Bcl-B protein and
CC   its corresponding nucleic acid. Bcl-B is useful in treating a subject
CC   having or at risk of a disorder associated with apoptosis, such as cell
CC   degenerative or proliferative disorder like neural or muscle
CC   degeneration, e.g. Alzheimer's disease, Creutzfeldt-Jacob's disease
CC   (CJD), Machado-Joseph disease (MJD), Parkinson's disease, Huntington's
CC   disease (HD), spinocerebellar ataxias 1, 2 and 6 (SCA-1, -2 and -6),
CC   dentatorubropallidoluysian atrophy (DRPLA), Kennedy's disease, stroke,
CC   ischaemia, head trauma, neoplasia, autoimmune disorder or fibrotic
CC   condition. The transgenic animals are used as in vivo models to study
CC   apoptosis and potential therapies for apoptosis. The present sequence is
CC   human Bcl-B DNA. Bcl-B gene is located at chromosome 15
XX
SQ   Sequence 887 BP; 171 A; 270 C; 269 G; 177 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 887;  DB 6;  Length 887;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  887;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGCCAAGAAAACCAGCGAAGGCCCGGCCCCCCAGCAGAGGCCGGACCATGGTTGACCA 60

Qy         61 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTTGCGGGAGCGCACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGC 120

Qy        121 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGACTACCTGGGGTACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCC 180

Qy        181 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGAGGCCGCCGTGCTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTT 240

Qy        241 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTC 300

Qy        301 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CGTGCTCTCCGACAGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGC 360

Qy        361 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGACGCTGCTGGAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCA 420

Qy        421 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCCGCGGCTAAAGGAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTT 480

Qy        481 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTGAGCTCGCGGCTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGA 540

Qy        541 TGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGT 600

Qy        601 CCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATT 660

Qy        661 ATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 ATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAG 720

Qy        721 ATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGA 780

Qy        781 ATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACT 840

Qy        841 CAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
              |||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887


RESULT 2
AAZ90049
ID   AAZ90049 standard; cDNA; 1168 BP.
XX
AC   AAZ90049;
XX
DT   09-MAY-2000  (first entry)
XX
DE   Hydrophobic domain containing protein clone HP02403 nucleotide sequence.
XX
KW   Hydrophobic domain; clone HP02403; nutritional supplement; SCID; HIV;
KW   cell proliferation; immune stimulant; immune deficiency; tumour; pain;
KW   rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
KW   myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
KW   anti-inflammatory; infection; bodily characteristic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200000506-A2.
XX
PD   06-JAN-2000.
XX
PF   18-JUN-1999;   99WO-JP003242.
XX
PR   26-JUN-1998;   98JP-00180008.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Kimura T;
XX
DR   WPI; 2000-160665/14.
DR   P-PSDB; AAY78802.
XX
PT   Novel human proteins having hydrophobic domains used for research and
PT   diagnostic purposes.
XX
PS   Claim 4; Page 92-94; 117pp; English.
XX
CC   This sequence represents the hydrophobic domain containing protein, clone
CC   HP00631 nucleotide sequence. The sequence is isolated from a human
CC   stomach cancer cell line. The HP02403 protein contains one putative
CC   transmembrane domain. The protein shows homology to the japanese quail
CC   apoptosis regualtor NR-13. The invention relates to human proteins with
CC   hydrophobic domains, the DNA and the cDNA encoding them. The
CC   polynucleotides and proteins are predicted to have biological activities
CC   which make them suitable for treating, preventing or ameliorating medical
CC   conditions in humans and animals. Suggested activities include
CC   nutritional activity (nutritional source or supplement); cytokine and
CC   cell proliferation/differentiation activity; immune stimulating (e.g. as
CC   vaccines) or suppressing activity (e.g. to treat various immune
CC   deficiencies such as SCIDS or HIV, connective tissue disease, systemic
CC   lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC   inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC   dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC   and autoimmune inflammatory eye disease, as well as asthma, allergies and
CC   organ transplantation); haematopoiesis regulating activity (e.g. in
CC   treatment of myeloid or lymphoid cell deficiencies); tissue growth
CC   activity (e.g. wound healing and tissue repair, ulcers, burns,
CC   periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
CC   activity; haemostatic and thrombolytic activity (e.g. treating
```

```
CC   haemophilias); receptor/ligand activity; anti-inflammatory activity; and
CC   tumour inhibition activity. The polynucleotides are also stated to be
CC   useful for gene therapy. Other activities include inhibiting infections
CC   caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
CC   malaria); effecting bodily characteristics such as, e.g. weight, colour,
CC   skin, effecting biorhythms or caricadic cycles; enhancing fertility;
CC   treatment of depression; treatment of pain; hormonal or endocrine
CC   activity. The polynucleotides may also be used for recombinant expression
CC   of the protein
XX
SQ   Sequence 1168 BP; 276 A; 312 C; 300 G; 280 T; 0 U; 0 Other;

  Query Match           86.0%;  Score 763;  DB 3;  Length 1168;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  813;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         74 ACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGG 133
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ACCACCATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGG 60

Qy        134 TACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTG 193
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TACTGCGCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTG 120

Qy        194 CTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTC 253
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCGCTCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTC 180

Qy        254 GGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGAC 313
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGAC 240

Qy        314 AGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTG 373
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCCCCGGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTG 300

Qy        374 GAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAG 433
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GAGAGAGGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAG 360

Qy        434 GAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGG 493
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGCAGGAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGG 420

Qy        494 CTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCAC 553
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTCATGGGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCAC 480

Qy        554 TTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTG 613
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCTTCAGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTG 540

Qy        614 TCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAA 673
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCATGCTTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAA 600

Qy        674 CTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACA 733
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTTAACCCGCTTCTACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACA 660

Qy        734 AGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGA 793
               |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        661 AGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTACCTGAATGCATACAAGGA 720

Qy        794 GTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGAC 853
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGAC 780

Qy        854 AAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
               |||||||||||||||||||||||||||||||||
Db        781 AAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 814
```

```
RESULT 3
AAZ90039
ID   AAZ90039 standard; cDNA; 582 BP.
XX
AC   AAZ90039;
XX
DT   09-MAY-2000  (first entry)
XX
DE   Hydrophobic domain containing protein clone HP02403 coding sequence.
XX
KW   Hydrophobic domain; clone HP02403; nutritional supplement; SCID; HIV;
KW   cell proliferation; immune stimulant; immune deficiency; tumour; pain;
KW   rheumatoid arthritis; insulin dependent diabetes mellitus; fertility;
KW   myasthenia gravis; haematopoiesis regulator; tissue growth; depression;
KW   anti-inflammatory; infection; bodily characteristic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200000506-A2.
XX
PD   06-JAN-2000.
XX
PF   18-JUN-1999;   99WO-JP003242.
XX
PR   26-JUN-1998;   98JP-00180008.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Kimura T;
XX
DR   WPI; 2000-160665/14.
DR   P-PSDB; AAY78802.
XX
PT   Novel human proteins having hydrophobic domains used for research and
PT   diagnostic purposes.
XX
PS   Claim 3; Page 84; 117pp; English.
XX
CC   This sequence represents the hydrophobic domain containing protein, clone
CC   HP00631 coding region. The sequence is isolated from a human stomach
CC   cancer cell line. The HP02403 protein contains one putative transmembrane
CC   domain. The protein shows homology to the japanese quail apoptosis
CC   regualtor NR-13. The invention relates to human proteins with hydrophobic
CC   domains, the DNA and the cDNA encoding them. The polynucleotides and
CC   proteins are predicted to have biological activities which make them
CC   suitable for treating, preventing or ameliorating medical conditions in
CC   humans and animals. Suggested activities include nutritional activity
CC   (nutritional source or supplement); cytokine and cell
CC   proliferation/differentiation activity; immune stimulating (e.g. as
CC   vaccines) or suppressing activity (e.g. to treat various immune
CC   deficiencies such as SCIDS or HIV, connective tissue disease, systemic
CC   lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC   inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC   dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC   and autoimmune inflammatory eye disease, as well as asthma, allergies and
CC   organ transplantation); haematopoiesis regulating activity (e.g. in
CC   treatment of myeloid or lymphoid cell deficiencies); tissue growth
CC   activity (e.g. wound healing and tissue repair, ulcers, burns,
CC   periodontal disease); activin/inhibin activity; chemotactic/chemokinetic
CC   activity; haemostatic and thrombolytic activity (e.g. treating
CC   haemophilias); receptor/ligand activity; anti-inflammatory activity; and
CC   tumour inhibition activity. The polynucleotides are also stated to be
CC   useful for gene therapy. Other activities include inhibiting infections
CC   caused by bacteria, fungi, viruses and other parasites (e.g. Hepatitis,
CC   malaria); effecting bodily characteristics such as, e.g. weight, colour,
CC   skin, effecting biorhythms or caricadic cycles; enhancing fertility;
CC   treatment of depression; treatment of pain; hormonal or endocrine
CC   activity. The polynucleotides may also be used for recombinant expression
CC   of the protein
XX
SQ   Sequence 582 BP; 85 A; 193 C; 191 G; 113 T; 0 U; 0 Other;
```

```
  Query Match             65.6%;  Score 582;  DB 3;  Length 582;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-277;
  Matches  582;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         80 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 139
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCCGACCCGCTGCGGGAGCGCACCGAGCTGTTGCTGGCCGACTACCTGGGGTACTGC 60

Qy        140 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 199
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCGGGAACCCGGCACCCCCGAGCCGGCGCCATCCACGCCCGAGGCCGCCGTGCTGCGC 120

Qy        200 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 259
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCCGCGGCCGCCAGGTTACGGCAGATTCACCGGTCCTTTTTCTCCGCCTACCTCGGCTAC 180

Qy        260 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 319
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCC 240

Qy        320 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 379
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGCCCCACCTGGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGA 300

Qy        380 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 439
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGCCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAG 360

Qy        440 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 499
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGGGCGACGTCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATG 420

Qy        500 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 559
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGCAGCACCGCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTC 480

Qy        560 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 619
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGGACCCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGC 540

Qy        620 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 661
               |||||||||||||||||||||||||||||||||||||||||
Db        541 TTGTTAACAACAGCCTTCATTTATCTCTGGACACGATTATTA 582
```

```
RESULT 4
AAH47022
ID   AAH47022 standard; cDNA; 522 BP.
XX
AC   AAH47022;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Human Bcl-2-like polypeptide encoding cDNA (clone HLIBE40).
XX
KW   Bcl-2-like polypeptide; autoimmune disorder; allergy; immunomodulatory;
KW   respiratory; cardiovascular; antiarthritic; immunostimulant; vaccine;
KW   immunosuppressive; antiinflammatory; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             44. .448
FT                   /*tag= a
FT                   /product= "Bcl-2-like polypeptide"
FT                   /note= "gene No. 2"
XX
PN   WO200157060-A1.
XX
PD   09-AUG-2001.
XX
PF   31-JAN-2001; 2001WO-US003080.
XX
PR   01-FEB-2000; 2000US-0179487P.
PR   07-FEB-2000; 2000US-0180697P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Duan DR,  Ni J;
XX
DR   WPI; 2001-476279/51.
DR   P-PSDB; AAB85666.
XX
PT   Nucleic acids encoding human Bcl-2-like polypeptides, useful for
PT   preventing, diagnosing and/or treating.
XX
PS   Claim 1; Page 276; 285pp; English.
XX
CC   The invention provides nucleic acid molecules (NAM1) encoding 4 human Bcl
CC   -2-like polypeptides (PEP1). The NAM1 and PEP1 may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate Bcl-2-like polypeptides' expression. The NAM1 may be used
CC   to produce the soluble Bcl-2-like polypeptides by standard recombinant
CC   methodolgy. The polypeptides may also be used as antigens in the
CC   production of antibodies against Bcl-2 and in assays to identify
CC   modulators of Bcl-2 expression and activity. The anti-Bcl-2 antibodies
CC   and antagonists may be used to down regulate expression and activity. The
CC   anti-PEP1 antibodies may also be used as diagnostic agents for detecting
CC   the presence of Bcl-2 polyPs in samples (e.g. by enzyme linked
CC   immunosorbant assay (ELISA)). Disorders that may be prevented, diagnosed
CC   and/or treated by the above methods include, immunodeficiencies (e.g. a
CC   gammaglobulinemia and B cell lymphoproliferative disorder), autoimmune
CC   disorders (e.g. rheumatoid arthritis and Grave's disease), allergic
CC   reactions, inflammations, respiratory diseases and cardiovascular
CC   disorders (a full list of disorders is given in the specification). The
CC   present sequence represents a human Bcl-2-like polypeptide encoding cDNA
XX
SQ   Sequence 522 BP; 102 A; 148 C; 154 G; 118 T; 0 U; 0 Other;
```

```
  Query Match             55.0%;  Score 488;  DB 4;  Length 522;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-231;
  Matches  488;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        270 GCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCT 329
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTTCGAGCTGGTGGCGCTGATGGCGGATTCCGTGCTCTCCGACAGCCCCGGCCCCACCT 60

Qy        330 GGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGG 389
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGGGCAGAGTGGTGACGCTCGTGACCTTCGCAGGGACGCTGCTGGAGAGAGGGCCGCTGG 120

Qy        390 TGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACG 449
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGAGCAGGAGGGCGACG 180

Qy        450 TCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACC 509
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCGCCCGGGACTGCCAGCGCCTGGTGGCCTTGCTGAGCTCGCGGCTCATGGGGCAGCACC 240

Qy        510 GCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCT 569
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCGCCTGGCTGCAGGCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCT 300

Qy        570 TTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAA 629
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAA 360

Qy        630 CAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACCCGCTTCT 689
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAGCCTTCATTTATCTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACCCGCTTCT 420

Qy        690 ACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAGGGAAAGC 749
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAGGGAAAGC 480

Qy        750 ACCTTCCC 757
```

```
                 ||||||||
Db        481 ACCTTCCC 488


RESULT 5
ABK41913
ID   ABK41913 standard; cDNA; 548 BP.
XX
AC   ABK41913;
XX
DT   21-MAY-2002  (first entry)
XX
DE   cDNA encoding novel human connective tissue related polypeptide #301.
XX
KW   Human; connective tissue related disorder; cancer; gene therapy;
KW   cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200155343-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001322.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
PR   08-SEP-2000; 2000US-0231244P.
PR   08-SEP-2000; 2000US-0231413P.
PR   08-SEP-2000; 2000US-0231414P.
PR   08-SEP-2000; 2000US-0232080P.
PR   08-SEP-2000; 2000US-0232081P.
PR   12-SEP-2000; 2000US-0231968P.
PR   14-SEP-2000; 2000US-0232397P.
PR   14-SEP-2000; 2000US-0232398P.
PR   14-SEP-2000; 2000US-0232399P.
PR   14-SEP-2000; 2000US-0232400P.
PR   14-SEP-2000; 2000US-0232401P.
PR   14-SEP-2000; 2000US-0233063P.
PR   14-SEP-2000; 2000US-0233064P.
PR   14-SEP-2000; 2000US-0233065P.
PR   21-SEP-2000; 2000US-0234223P.
PR   21-SEP-2000; 2000US-0234274P.
PR   25-SEP-2000; 2000US-0234997P.
PR   25-SEP-2000; 2000US-0234998P.
PR   26-SEP-2000; 2000US-0235484P.
PR   27-SEP-2000; 2000US-0235834P.
PR   27-SEP-2000; 2000US-0235836P.
PR   29-SEP-2000; 2000US-0236327P.
PR   29-SEP-2000; 2000US-0236367P.
PR   29-SEP-2000; 2000US-0236368P.
PR   29-SEP-2000; 2000US-0236369P.
PR   29-SEP-2000; 2000US-0236370P.
PR   02-OCT-2000; 2000US-0236802P.
PR   02-OCT-2000; 2000US-0237037P.
PR   02-OCT-2000; 2000US-0237038P.
PR   02-OCT-2000; 2000US-0237039P.
PR   02-OCT-2000; 2000US-0237040P.
PR   13-OCT-2000; 2000US-0239935P.
PR   13-OCT-2000; 2000US-0239937P.
PR   20-OCT-2000; 2000US-0240960P.
PR   20-OCT-2000; 2000US-0241221P.
PR   20-OCT-2000; 2000US-0241785P.
PR   20-OCT-2000; 2000US-0241786P.
PR   20-OCT-2000; 2000US-0241787P.
PR   20-OCT-2000; 2000US-0241808P.
PR   20-OCT-2000; 2000US-0241809P.
PR   20-OCT-2000; 2000US-0241826P.
PR   01-NOV-2000; 2000US-0244617P.
PR   08-NOV-2000; 2000US-0246474P.
PR   08-NOV-2000; 2000US-0246475P.
PR   08-NOV-2000; 2000US-0246476P.
PR   08-NOV-2000; 2000US-0246477P.
PR   08-NOV-2000; 2000US-0246478P.
PR   08-NOV-2000; 2000US-0246523P.
PR   08-NOV-2000; 2000US-0246524P.
PR   08-NOV-2000; 2000US-0246525P.
PR   08-NOV-2000; 2000US-0246526P.
PR   08-NOV-2000; 2000US-0246527P.
PR   08-NOV-2000; 2000US-0246528P.
PR   08-NOV-2000; 2000US-0246532P.
PR   08-NOV-2000; 2000US-0246609P.
PR   08-NOV-2000; 2000US-0246610P.
PR   08-NOV-2000; 2000US-0246611P.
PR   08-NOV-2000; 2000US-0246613P.
PR   17-NOV-2000; 2000US-0249207P.
PR   17-NOV-2000; 2000US-0249208P.
PR   17-NOV-2000; 2000US-0249209P.
PR   17-NOV-2000; 2000US-0249210P.
PR   17-NOV-2000; 2000US-0249211P.
PR   17-NOV-2000; 2000US-0249212P.
PR   17-NOV-2000; 2000US-0249213P.
PR   17-NOV-2000; 2000US-0249214P.
PR   17-NOV-2000; 2000US-0249215P.
PR   17-NOV-2000; 2000US-0249216P.
PR   17-NOV-2000; 2000US-0249217P.
PR   17-NOV-2000; 2000US-0249218P.
PR   17-NOV-2000; 2000US-0249244P.
PR   17-NOV-2000; 2000US-0249245P.
PR   17-NOV-2000; 2000US-0249264P.
PR   17-NOV-2000; 2000US-0249265P.
PR   17-NOV-2000; 2000US-0249297P.
```

```
PR   17-NOV-2000; 2000US-0249299P.
PR   17-NOV-2000; 2000US-0249300P.
PR   01-DEC-2000; 2000US-0250160P.
PR   01-DEC-2000; 2000US-0250391P.
PR   05-DEC-2000; 2000US-0251030P.
PR   05-DEC-2000; 2000US-0251988P.
PR   05-DEC-2000; 2000US-0256719P.
PR   06-DEC-2000; 2000US-0251479P.
PR   08-DEC-2000; 2000US-0251856P.
PR   08-DEC-2000; 2000US-0251868P.
PR   08-DEC-2000; 2000US-0251869P.
PR   08-DEC-2000; 2000US-0251989P.
PR   08-DEC-2000; 2000US-0251990P.
PR   11-DEC-2000; 2000US-0254097P.
PR   05-JAN-2001; 2001US-0259678P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-565190/63.
DR   P-PSDB; AAU86735.
XX
PT   Nucleic acid encoding novel connective tissue associated polypeptides,
PT   used in diagnosing, preventing, treating or ameliorating a disorder such
PT   as cancer or rheumatoid arthritis.
XX
PS   Claim 4; SEQ ID NO 311; 673pp; English.
XX
CC   The present invention relates to the isolation of novel human connective
CC   tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
CC   (cDNA and genomic) sequences encoding them. The sequences of the
CC   invention are useful in the diagnosis, treatment, prevention and/or
CC   prognosis of diseases associated with connective tissue(s), including
CC   cancer. The polynucleotide sequences of the invention are also useful in
CC   gene therapy. ABK41613-ABK42101 represent cDNA sequences encoding the
CC   novel human connective tissue related polypeptides. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 548 BP; 99 A; 154 C; 167 G; 121 T; 0 U; 7 Other;

  Query Match             25.4%;  Score 225;  DB 4;  Length 548;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-101;
  Matches  225;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        524 GCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTT 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 GCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTT 383

Qy        584 TGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTAT 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 TGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTAT 443

Qy        644 CTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGT 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGT 503

Qy        704 GACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAGGGAAAG 748
              |||||||||||||||||||||||||||||||||||||||||||||
Db        504 GACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAGGGAAAG 548


RESULT 6
ADB59580
ID   ADB59580 standard; cDNA; 548 BP.
XX
AC   ADB59580;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Connective tissue related polynucleotide #301.
XX
KW   cytostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;
KW   antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;
KW   antiinflammatory; antiallergic; antiasthmatic; dermatological;
KW   nephrotopic; virucide; fungicide; antibacterial; antiparasitic;
KW   gene therapy; ds; connective tissues disorder; rheumatoid arthritis;
KW   systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;
KW   cancer metastasis; neoplasia; leukaemia; neurodegenerative disorder;
KW   Alzheimer's disease; Parkinson's disease; cardiovascular disease;
KW   atherosclerosis; myocarditis; cardiopulmonary bypass complication;
KW   autoimmune disease; multiple sclerosis; allergic reaction; asthma;
KW   rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;
KW   gastrointestinal disorder; inflammatory bowel disease;
KW   organ transplant rejection; immune system disorder; Bruton's disease;
KW   X-linked lymphoproliferative syndrome;
KW   B-cell lymphoproliferative disorder; HIV; AIDS; infection;
KW   chromosome identification; chromosome mapping;
KW   connective tissue related polynucleotide; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003054375-A1.
XX
PD   20-MAR-2003.
XX
PF   07-MAR-2002; 2002US-00092154.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
PR   16-MAR-2000; 2000US-0189874P.
PR   17-MAR-2000; 2000US-0190076P.
PR   18-APR-2000; 2000US-0198123P.
PR   19-MAY-2000; 2000US-0205515P.
PR   07-JUN-2000; 2000US-0209467P.
PR   28-JUN-2000; 2000US-0214886P.
PR   30-JUN-2000; 2000US-0215135P.
PR   07-JUL-2000; 2000US-0216647P.
PR   07-JUL-2000; 2000US-0216880P.
PR   11-JUL-2000; 2000US-0217487P.
PR   11-JUL-2000; 2000US-0217496P.
PR   14-JUL-2000; 2000US-0218290P.
PR   26-JUL-2000; 2000US-0220963P.
PR   26-JUL-2000; 2000US-0220964P.
PR   14-AUG-2000; 2000US-0224518P.
PR   14-AUG-2000; 2000US-0224519P.
PR   14-AUG-2000; 2000US-0225213P.
PR   14-AUG-2000; 2000US-0225214P.
PR   14-AUG-2000; 2000US-0225266P.
PR   14-AUG-2000; 2000US-0225267P.
PR   14-AUG-2000; 2000US-0225268P.
PR   14-AUG-2000; 2000US-0225270P.
PR   14-AUG-2000; 2000US-0225447P.
PR   14-AUG-2000; 2000US-0225757P.
PR   14-AUG-2000; 2000US-0225758P.
PR   14-AUG-2000; 2000US-0225759P.
PR   18-AUG-2000; 2000US-0226279P.
PR   22-AUG-2000; 2000US-0226681P.
PR   22-AUG-2000; 2000US-0226868P.
PR   22-AUG-2000; 2000US-0227182P.
PR   23-AUG-2000; 2000US-0227009P.
PR   30-AUG-2000; 2000US-0228924P.
PR   01-SEP-2000; 2000US-0229287P.
PR   01-SEP-2000; 2000US-0229343P.
PR   01-SEP-2000; 2000US-0229344P.
PR   01-SEP-2000; 2000US-0229345P.
PR   05-SEP-2000; 2000US-0229509P.
PR   05-SEP-2000; 2000US-0229513P.
PR   06-SEP-2000; 2000US-0230437P.
PR   06-SEP-2000; 2000US-0230438P.
PR   08-SEP-2000; 2000US-0231242P.
PR   08-SEP-2000; 2000US-0231243P.
```

PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764847.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI; 2003-634869/60.
DR P-PSDB; ADB60069.
XX
PT New connective tissue-related polypeptides and polynucleotides, useful
PT for treating, preventing and/or prognosing e.g. disorders of connective
PT tissue, (e.g. rheumatoid arthritis), cancers, cancer metastases and/or
PT neoplasias.
XX
PS Claim 1; SEQ ID NO 311; 248pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I), which
CC comprises a sequence that is at least 95 % identical to a connective
CC tissue-related polynucleotide encoding connective tissue antigens (CTA).
CC The polypeptide or polynucleotide is useful for preventing, treating, or
CC ameliorating medical conditions in a mammal. The connective tissue
CC polypeptides, polynucleotides and antibodies are particularly useful for
CC treating, preventing and/or prognosing disorders of connective tissues
CC (e.g. rheumatoid arthritis, discoid and systemic lupus erythematosus,
CC scleroderma, or Sjogren's syndrome), cancers, cancer metastases and/or
CC neoplasias (e.g. leukaemia), neurodegenerative disorders (e.g.
CC Alzheimer's disease, or Parkinson's disease), cardiovascular diseases
CC (e.g. atherosclerosis, myocarditis or cardiopulmonary bypass
CC complications), autoimmune diseases (e.g. systemic lupus erythematosus,
CC rheumatoid arthritis, or multiple sclerosis), allergic reactions (e.g.

Query Match 25.4%; Score 225; DB 9; Length 548;
Best Local Similarity 100.0%; Pred. No. 9.2e-101;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        524 GCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTT 583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 GCTCAGGGCGGCTGGGATGGCTTTTGTCACTTCTTCAGGACCCCCTTTCCACTGGCTTTT 383

Qy        584 TGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTAT 643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 TGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTTAACAACAGCCTTCATTTAT 443

Qy        644 CTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGT 703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CTCTGGACACGATTATTATGAGTTTTAAAACTTTTAACCCGCTTCTACCTGCCCAACTGT 503

Qy        704 GACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAGGGAAAG 748
              |||||||||||||||||||||||||||||||||||||||||||||
Db        504 GACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAGGGAAAG 548
```

RESULT 7
ABN42801
ID ABN42801 standard; DNA; 60 BP.
XX

```
AC   ABN42801;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:15549.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB001903.
XX
PR   28-JUL-2000; 2000US-0221607P.
PR   02-MAY-2001; 2001US-0287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of a
PT   genome, useful for detecting tissue-, pathology-, and developmental-
PT   specific genes.
XX
PS   Example 1; SEQ ID NO 15549; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the (sub-
CC   )transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises several
CC   oligonucleotides, each capable of hybridising selectively to a set of
CC   messenger RNAs transcribed from a given transcription unit of the genome,
CC   which encodes one or more messenger RNA splice variants. The
CC   oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a particular
CC   biological or pathological state, and so allowing the detection of tissue
CC   - and pathology-specific genes such as those genes only expressed in
CC   specific tissue under a specific pathological condition; to detect
CC   developmental specific genes; and to detect RNA transcripts and splice
CC   variants of a transcriptome of a patient suffering from a particular
CC   disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC   rats, humans and mice, which are used in the exemplification of the
CC   present invention. N.B. The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 60 BP; 9 A; 16 C; 13 G; 22 T; 0 U; 0 Other;

  Query Match             6.8%;  Score 60;  DB 6;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-19;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        565 CCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTT 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCCTTTCCACTGGCTTTTTGGAGAAAACAGCTGGTCCAGGCTTTTCTGTCATGCTTGTT 60


RESULT 8
AAD46685/c
ID   AAD46685 standard; DNA; 24 BP.
XX
AC   AAD46685;
XX
DT   27-JAN-2003  (first entry)
XX
DE   Human Bcl-B cDNA cloning reverse RT-PCR primer, NK0121.
XX
KW   Human; Bcl-2; Bcl-B; therapy; apoptosis; cell degenerative disorder;
KW   proliferative disorder; muscle degeneration; Alzheimer's disease; CJD;
KW   Creutzfeldt-Jacob's disease; Machado-Joseph disease; MJD; transgenic;
KW   Parkinson's disease; Huntington's disease; HD; spinocerebellar ataxia;
KW   SCA; dentatorubropallidoluysian atrophy; DRPLA; Kennedy's disease;
KW   stroke; ischaemia; head trauma; neoplasia; anticonvulsant; vulnerary;
KW   nootropic; neuroprotective; cytostatic; immunosuppressive; vasotropic;
KW   cerebroprotective; autoimmune disorder; reverse transcription; RT; PCR;
KW   primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200272601-A2.
XX
PD   19-SEP-2002.
XX
PF   07-FEB-2002; 2002WO-US003547.
XX
PR   07-FEB-2001; 2001US-0267166P.
PR   07-FEB-2002; 2002US-00071174.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Reed JC,  Ke N,  Godzik A;
XX
DR   WPI; 2002-723312/78.
XX
PT   New isolated or recombinant Bcl-B nucleic acids and polypeptides, for
PT   treating a disorder associated with apoptosis, such as cell degenerative
PT   or proliferative disorder e.g. cancer, Alzheimer's disease or Parkinson's
PT   disease.
XX
PS   Example 1; Page 47; 82pp; English.
XX
CC   The invention relates to human member of Bcl-2 family Bcl-B protein and
CC   its corresponding nucleic acid. Bcl-B is useful in treating a subject
CC   having or at risk of a disorder associated with apoptosis, such as cell
CC   degenerative or proliferative disorder like neural or muscle
CC   degeneration, e.g. Alzheimer's disease, Creutzfeldt-Jacob's disease
CC   (CJD), Machado-Joseph disease (MJD), Parkinson's disease, Huntington's
CC   disease (HD), spinocerebellar ataxias 1, 2 and 6 (SCA-1, -2 and -6),
CC   dentatorubropallidoluysian atrophy (DRPLA), Kennedy's disease, stroke,
CC   ischaemia, head trauma, neoplasia, autoimmune disorder or fibrotic
CC   condition. The transgenic animals are used as in vivo models to study
CC   apoptosis and potential therapies for apoptosis. The present sequence is
CC   a reverse transcription (RT)-PCR primer used for cloning human Bcl-B cDNA
XX
SQ   Sequence 24 BP; 8 A; 6 C; 5 G; 5 T; 0 U; 0 Other;

  Query Match             2.7%;  Score 24;  DB 6;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 0.3;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        858 ATGCAAATGGCTCTTCCTTGAGTG 881
              ||||||||||||||||||||||||
Db         24 ATGCAAATGGCTCTTCCTTGAGTG 1


RESULT 9
AAD46684
ID   AAD46684 standard; DNA; 23 BP.
XX
AC   AAD46684;
XX
DT   27-JAN-2003  (first entry)
XX
DE   Human Bcl-B cDNA cloning forward RT-PCR primer, NK01.
XX
```

KW Human; Bcl-2; Bcl-B; therapy; apoptosis; cell degenerative disorder;
KW proliferative disorder; muscle degeneration; Alzheimer's disease; CJD;
KW Creutzfeldt-Jacob's disease; Machado-Joseph disease; MJD; transgenic;
KW Parkinson's disease; Huntington's disease; HD; spinocerebellar ataxia;
KW SCA; dentatorubropallidoluysian atrophy; DRPLA; Kennedy's disease;
KW stroke; ischaemia; head trauma; neoplasia; anticonvulsant; vulnerary;
KW nootropic; neuroprotective; cytostatic; immunosuppressive; vasotropic;
KW cerebroprotective; autoimmune disorder; reverse transcription; RT; PCR;
KW primer; ss.
XX
OS Homo sapiens.
XX
PN WO200272601-A2.
XX
PD 19-SEP-2002.
XX
PF 07-FEB-2002; 2002WO-US003547.
XX
PR 07-FEB-2001; 2001US-0267166P.
PR 07-FEB-2002; 2002US-00071174.
XX
PA (BURN-) BURNHAM INST.
XX
PI Reed JC, Ke N, Godzik A;
XX
DR WPI; 2002-723312/78.
XX
PT New isolated or recombinant Bcl-B nucleic acids and polypeptides, for
PT treating a disorder associated with apoptosis, such as cell degenerative
PT or proliferative disorder e.g. cancer, Alzheimer's disease or Parkinson's
PT disease.
XX
PS Example 1; Page 47; 82pp; English.
XX
CC The invention relates to human member of Bcl-2 family Bcl-B protein and
CC its corresponding nucleic acid. Bcl-B is useful in treating a subject
CC having or at risk of a disorder associated with apoptosis, such as cell
CC degenerative or proliferative disorder like neural or muscle
CC degeneration, e.g. Alzheimer's disease, Creutzfeldt-Jacob's disease
CC (CJD), Machado-Joseph disease (MJD), Parkinson's disease, Huntington's
CC disease (HD), spinocerebellar ataxias 1, 2 and 6 (SCA-1, -2 and -6),
CC dentatorubropallidoluysian atrophy (DRPLA), Kennedy's disease, stroke,
CC ischaemia, head trauma, neoplasia, autoimmune disorder or fibrotic
CC condition. The transgenic animals are used as in vivo models to study
CC apoptosis and potential therapies for apoptosis. The present sequence is
CC a reverse transcription (RT)-PCR primer used for cloning human Bcl-B cDNA
XX
SQ Sequence 23 BP; 9 A; 6 C; 8 G; 0 T; 0 U; 0 Other;

Query Match 2.6%; Score 23; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCCAAGAAAACCAGCGAAGG 23
||||||||||||||||||||||||
Db 1 CGGGCCAAGAAAACCAGCGAAGG 23

RESULT 10
AAD46689
ID AAD46689 standard; DNA; 22 BP.
XX
AC AAD46689;
XX
DT 27-JAN-2003 (first entry)
XX
DE Human Bcl-B cDNA amplifying forward RT-PCR primer, NK0120.
XX
KW Human; Bcl-2; Bcl-B; therapy; apoptosis; cell degenerative disorder;
KW proliferative disorder; muscle degeneration; Alzheimer's disease; CJD;
KW Creutzfeldt-Jacob's disease; Machado-Joseph disease; MJD; transgenic;
KW Parkinson's disease; Huntington's disease; HD; spinocerebellar ataxia;
KW SCA; dentatorubropallidoluysian atrophy; DRPLA; Kennedy's disease;
KW stroke; ischaemia; head trauma; neoplasia; anticonvulsant; vulnerary;
KW nootropic; neuroprotective; cytostatic; immunosuppressive; vasotropic;
KW cerebroprotective; autoimmune disorder; reverse transcription; RT; PCR;
KW primer; ss.
XX
OS Homo sapiens.
XX
PN WO200272601-A2.
XX
PD 19-SEP-2002.
XX
PF 07-FEB-2002; 2002WO-US003547.
XX
PR 07-FEB-2001; 2001US-0267166P.
PR 07-FEB-2002; 2002US-00071174.
XX
PA (BURN-) BURNHAM INST.
XX
PI Reed JC, Ke N, Godzik A;
XX
DR WPI; 2002-723312/78.
XX
PT New isolated or recombinant Bcl-B nucleic acids and polypeptides, for
PT treating a disorder associated with apoptosis, such as cell degenerative
PT or proliferative disorder e.g. cancer, Alzheimer's disease or Parkinson's
PT disease.
XX
PS Example 1; Page 48; 82pp; English.
XX
CC The invention relates to human member of Bcl-2 family Bcl-B protein and
CC its corresponding nucleic acid. Bcl-B is useful in treating a subject
CC having or at risk of a disorder associated with apoptosis, such as cell
CC degenerative or proliferative disorder like neural or muscle
CC degeneration, e.g. Alzheimer's disease, Creutzfeldt-Jacob's disease
CC (CJD), Machado-Joseph disease (MJD), Parkinson's disease, Huntington's
CC disease (HD), spinocerebellar ataxias 1, 2 and 6 (SCA-1, -2 and -6),
CC dentatorubropallidoluysian atrophy (DRPLA), Kennedy's disease, stroke,
CC ischaemia, head trauma, neoplasia, autoimmune disorder or fibrotic
CC condition. The transgenic animals are used as in vivo models to study
CC apoptosis and potential therapies for apoptosis. The present sequence is
CC a reverse transcription (RT)-PCR primer used for amplifying human Bcl-B
CC cDNA
XX
SQ Sequence 22 BP; 2 A; 6 C; 8 G; 6 T; 0 U; 0 Other;

Query Match 2.5%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 338 GTGGTGACGCTCGTGACCTTCG 359
||||||||||||||||||||||
Db 1 GTGGTGACGCTCGTGACCTTCG 22

RESULT 11
AAD46686
ID AAD46686 standard; DNA; 28 BP.
XX
AC AAD46686;
XX
DT 27-JAN-2003 (first entry)
XX
DE Human Bcl-B DNA amplifying forward PCR primer, NK0101.
XX
KW Human; Bcl-2; Bcl-B; therapy; apoptosis; cell degenerative disorder;
KW proliferative disorder; muscle degeneration; Alzheimer's disease; CJD;
KW Creutzfeldt-Jacob's disease; Machado-Joseph disease; MJD; transgenic;
KW Parkinson's disease; Huntington's disease; HD; spinocerebellar ataxia;
KW SCA; dentatorubropallidoluysian atrophy; DRPLA; Kennedy's disease;
KW stroke; ischaemia; head trauma; neoplasia; anticonvulsant; vulnerary;
KW nootropic; neuroprotective; cytostatic; immunosuppressive; vasotropic;
KW cerebroprotective; autoimmune disorder; PCR; primer; ss.
XX

OS Homo sapiens.
XX
PN WO200272601-A2.
XX
PD 19-SEP-2002.
XX
PF 07-FEB-2002; 2002WO-US003547.
XX
PR 07-FEB-2001; 2001US-0267166P.
PR 07-FEB-2002; 2002US-00071174.
XX
PA (BURN-) BURNHAM INST.
XX
PI Reed JC, Ke N, Godzik A;
XX
DR WPI; 2002-723312/78.
XX
PT New isolated or recombinant Bcl-B nucleic acids and polypeptides, for
PT treating a disorder associated with apoptosis, such as cell degenerative
PT or proliferative disorder e.g. cancer, Alzheimer's disease or Parkinson's
PT disease.
XX
PS Example 1; Page 47; 82pp; English.
XX
CC The invention relates to human member of Bcl-2 family Bcl-B protein and
CC its corresponding nucleic acid. Bcl-B is useful in treating a subject
CC having or at risk of a disorder associated with apoptosis, such as cell
CC degenerative or proliferative disorder like neural or muscle
CC degeneration, e.g. Alzheimer's disease, Creutzfeldt-Jacob's disease
CC (CJD), Machado-Joseph disease (MJD), Parkinson's disease, Huntington's
CC disease (HD), spinocerebellar ataxias 1, 2 and 6 (SCA-1, -2 and -6),
CC dentatorubropallidoluysian atrophy (DRPLA), Kennedy's disease, stroke,
CC ischaemia, head trauma, neoplasia, autoimmune disorder or fibrotic
CC condition. The transgenic animals are used as in vivo models to study
CC apoptosis and potential therapies for apoptosis. The present sequence is
CC a PCR primer used for amplifying human Bcl-B DNA
XX
SQ Sequence 28 BP; 6 A; 4 C; 11 G; 7 T; 0 U; 0 Other;

Query Match 2.5%; Score 22; DB 6; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 CATGGTTGACCAGTTGCGGGAG 70
||||||||||||||||||||||
Db 7 CATGGTTGACCAGTTGCGGGAG 28

RESULT 12
AAD46688/c
ID AAD46688 standard; DNA; 34 BP.
XX
AC AAD46688;
XX
DT 27-JAN-2003 (first entry)
XX
DE Human Bcl-B mutagenic PCR primer, NK0131.
XX
KW Human; Bcl-2; Bcl-B; therapy; apoptosis; cell degenerative disorder;
KW proliferative disorder; muscle degeneration; Alzheimer's disease; CJD;
KW Creutzfeldt-Jacob's disease; Machado-Joseph disease; MJD; transgenic;
KW Parkinson's disease; Huntington's disease; HD; spinocerebellar ataxia;
KW SCA; dentatorubropallidoluysian atrophy; DRPLA; Kennedy's disease;
KW stroke; ischaemia; head trauma; neoplasia; anticonvulsant; vulnerary;
KW nootropic; neuroprotective; cytostatic; immunosuppressive; vasotropic;
KW cerebroprotective; autoimmune disorder; PCR; primer; ss.
XX
OS Homo sapiens.
XX
PN WO200272601-A2.
XX
PD 19-SEP-2002.
XX
PF 07-FEB-2002; 2002WO-US003547.
XX
PR 07-FEB-2001; 2001US-0267166P.
PR 07-FEB-2002; 2002US-00071174.
XX
PA (BURN-) BURNHAM INST.
XX
PI Reed JC, Ke N, Godzik A;
XX
DR WPI; 2002-723312/78.
XX
PT New isolated or recombinant Bcl-B nucleic acids and polypeptides, for
PT treating a disorder associated with apoptosis, such as cell degenerative
PT or proliferative disorder e.g. cancer, Alzheimer's disease or Parkinson's
PT disease.
XX
PS Example 1; Page 47; 82pp; English.
XX
CC The invention relates to human member of Bcl-2 family Bcl-B protein and
CC its corresponding nucleic acid. Bcl-B is useful in treating a subject
CC having or at risk of a disorder associated with apoptosis, such as cell
CC degenerative or proliferative disorder like neural or muscle
CC degeneration, e.g. Alzheimer's disease, Creutzfeldt-Jacob's disease
CC (CJD), Machado-Joseph disease (MJD), Parkinson's disease, Huntington's
CC disease (HD), spinocerebellar ataxias 1, 2 and 6 (SCA-1, -2 and -6),
CC dentatorubropallidoluysian atrophy (DRPLA), Kennedy's disease, stroke,
CC ischaemia, head trauma, neoplasia, autoimmune disorder or fibrotic
CC condition. The transgenic animals are used as in vivo models to study
CC apoptosis and potential therapies for apoptosis. The present sequence is
CC a PCR primer used for constructing human Bcl-B mutant DNA
XX
SQ Sequence 34 BP; 8 A; 10 C; 7 G; 9 T; 0 U; 0 Other;

Query Match 2.5%; Score 22; DB 6; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 573 CACTGGCTTTTTGGAGAAAACA 594
||||||||||||||||||||||
Db 34 CACTGGCTTTTTGGAGAAAACA 13

RESULT 13
ABQ44403
ID ABQ44403 standard; DNA; 874 BP.
XX
AC ABQ44403;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30994.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.

XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 874 BP; 374 A; 291 C; 90 G; 119 T; 0 U; 0 Other;

Query Match 2.5%; Score 22; DB 6; Length 874;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 TCCTTTTTCTCCGCCTACCTCG 254
||||||||||||||||||||||
Db 332 TCCTTTTTCTCCGCCTACCTCG 353

RESULT 14
ABQ44402/c
ID ABQ44402 standard; DNA; 874 BP.
XX
AC ABQ44402;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 30993.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 874 BP; 119 A; 90 C; 291 G; 374 T; 0 U; 0 Other;

Query Match 2.5%; Score 22; DB 6; Length 874;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 233 TCCTTTTTCTCCGCCTACCTCG 254
||||||||||||||||||||||
Db 543 TCCTTTTTCTCCGCCTACCTCG 522

RESULT 15
ABN83959/c
ID ABN83959 standard; DNA; 5147 BP.
XX
AC ABN83959;
XX
DT 06-SEP-2002 (first entry)
XX
DE Human gene sequence #6.
XX
KW Human; brain; tonsil; hippocampus; foetal brain; diagnosis; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 2074. .2349
FT /*tag= a
FT /partial
FT /note= "no start codon present"
XX
PN WO200252005-A1.
XX
PD 04-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-JP011217.
XX
PR 22-DEC-2000; 2000JP-00389742.
XX
PA (KAZU-) KAZUSA DNA RES INST FOUND.
PA (CELE-) CELESTAR LEXICO-SCI LTD.
XX
PI Ohara O, Nagase T, Nakajima D;
XX
DR WPI; 2002-500762/53.
DR P-PSDB; ABB97939.
XX
PT Genes and their expression products cloned from human cDNA libraries for

PT treatment and diagnosis of diseases associated with their expression.
XX
PS Claim 1(a); Page 63-66; 238pp; Japanese.
XX
CC The invention relates to DNA encoding polypeptides directly cloned from
CC cDNA libraries originating in adult whole brain, human tonsil, human
CC adult hippocampus and human foetal whole brain. Polypeptides and
CC polynucleotides of the invention may be used in the investigation of
CC differential expression of the DNA sequences in normal subjects and
CC disease patients. They may also be used in the production of antibodies,
CC oligonucleotide probes and DNA chips for diagnosis and identification of
CC drugs for treatment of diseases with which the DNA sequences are
CC associated. The sequences given in records ABN83954-ABN83984 represent
CC human gene sequences of the invention
XX
SQ Sequence 5147 BP; 1212 A; 1455 C; 1334 G; 1146 T; 0 U; 0 Other;

Query Match 2.5%; Score 22; DB 6; Length 5147;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 512 GCCTGGCTGCAGGCTCAGGGCG 533
||||||||||||||||||||||
Db 4330 GCCTGGCTGCAGGCTCAGGGCG 4309

Search completed: November 7, 2004, 03:03:12
Job time : 688 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: November 7, 2004, 02:37:47 ; Search time 4548 Seconds
(without alignments)
7106.869 Million cell updates/sec

Title: US-10-071-174-1
Perfect score: 887
Sequence: 1 cgggccaagaaaaccagcga.....ctcttccttgagtgaaagaa 887

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 32822875 seqs, 18219865908 residues

Word size : 0

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	197	22.2	715	6	CA777633	ip19a09.y
c	2	167	18.8	535	5	BX119239	BX119239 BX119239
c	3	111	12.5	493	1	AI813346	AI813346 wj33g03.x
	4	91	10.3	130	1	AA098865	AA098865 zk84f02.s
	5	53	6.0	206	1	AA005293	AA005293 zh93a11.r
c	6	30	3.4	478	7	R53538	R53538 yg84c01.s1
c	7	27	3.0	939	9	CNS02C0P	AL190546 Tetraodon
	8	22	2.5	316	7	R85550	R85550 yo38d07.s1
	9	22	2.5	601	8	BZ869242	BZ869242 CH240_211
c	10	21	2.4	441	8	AZ617122	AZ617122 1M0448D17
c	11	21	2.4	462	5	BY590905	BY590905 BY590905
	12	21	2.4	479	8	AQ294438	AQ294438 HS_3014_B
	13	21	2.4	494	1	AA823641	AA823641 vr69a10.s
	14	21	2.4	657	2	BB628603	BB628603 BB628603
c	15	21	2.4	803	9	CNS0413S	AL269713 Tetraodon
	16	21	2.4	881	2	BF532088	BF532088 602073207
	17	21	2.4	931	4	BG173025	BG173025 602335506
	18	21	2.4	2780	3	AK035932	AK035932 Mus muscu
	19	20	2.3	352	1	AA426934	AA426934 vf22b11.s
c	20	20	2.3	415	5	BY059827	BY059827 BY059827
	21	20	2.3	448	6	CA560084	CA560084 K0267E03-
c	22	20	2.3	454	1	AJ631643	AJ631643 AJ631643
	23	20	2.3	455	8	AQ457939	AQ457939 HS_5189_B
	24	20	2.3	478	6	CA561864	CA561864 K0293B04-
	25	20	2.3	524	6	CA560564	CA560564 K0274C06-
c	26	20	2.3	545	4	BG071824	BG071824 H3103E12-
	27	20	2.3	588	6	CA557550	CA557550 K0229B10-
	28	20	2.3	596	7	CO702056	CO702056 DG32-220o
	29	20	2.3	613	5	BP464181	BP464181 BP464181
	30	20	2.3	626	4	BG080862	BG080862 H3058C03-
	31	20	2.3	649	7	CF915355	CF915355 B0978E08-
	32	20	2.3	710	8	BZ006250	BZ006250 oei81e02.
	33	20	2.3	728	9	BX969107	BX969107 Forward s
	34	20	2.3	736	6	BY735639	BY735639 BY735639
	35	20	2.3	758	7	CO799635	CO799635 AGENCOURT
	36	20	2.3	763	7	CO808612	CO808612 AGENCOURT
	37	20	2.3	772	7	CO814648	CO814648 AGENCOURT
	38	20	2.3	783	8	BZ452477	BZ452477 BONPM15TR
	39	20	2.3	791	5	BU941870	BU941870 AGENCOURT
	40	20	2.3	800	9	CL251188	CL251188 ZMMBBb060
	41	20	2.3	821	7	CO806531	CO806531 AGENCOURT
	42	20	2.3	852	8	BZ567066	BZ567066 pacs2-164
	43	20	2.3	875	7	CO798715	CO798715 AGENCOURT
	44	20	2.3	893	7	CO811958	CO811958 AGENCOURT
	45	19	2.1	190	4	BI301953	BI301953 UI-R-DL0-

ALIGNMENTS

RESULT 1
CA777633/c
LOCUS CA777633 715 bp mRNA linear EST 03-DEC-2002
DEFINITION ip19a09.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6217625 5' similar to TR:Q9Z0F3 Q9Z0F3 BCL-2 HOMOLOG. ;, mRNA sequence.
ACCESSION CA777633
VERSION CA777633.1 GI:26015508
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 715)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 392.
FEATURES Location/Qualifiers
source 1. .715
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6217625"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming.

Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 22.2%; Score 197; DB 6; Length 715;
Best Local Similarity 100.0%; Pred. No. 5.6e-95;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        691 CCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAGGGAAAGCA 750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 CCTGCCCAACTGTGACCAACTAAATGACAGATGTGTGAGAACAAGAACTGAGGGAAAGCA 522

Qy        751 CCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGAGTCCTGAGGTGGTGATT 810
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 CCTTCCCCCACCCCAGACGTTTTTATCTGAATGCATACAAGGAGTCCTGAGGTGGTGATT 462

Qy        811 TGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGACAAGAATGCAAATGGCTC 870
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 TGGCCAGTGTTTTAACTTGTGACAAGTACTCAGGTGTGAGGACAAGAATGCAAATGGCTC 402

Qy        871 TTCCTTGAGTGAAAGAA 887
              |||||||||||||||||
Db        401 TTCCTTGAGTGAAAGAA 385
```

RESULT 2
BX119239/c
LOCUS BX119239 535 bp mRNA linear EST 10-FEB-2003
DEFINITION BX119239 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGp998N01170 ; IMAGE:40052, mRNA sequence.
ACCESSION BX119239
VERSION BX119239.1 GI:27841652
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 535)
AUTHORS Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M., Radelof,U., Schneider,D. and Korn,B.
TITLE Human UnigeneSet - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998N01170.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer: M13u, Primer sequence: CGTTGTAAAACGACGGCCAGT.
FEATURES Location/Qualifiers
source 1. .535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998N01170 ; IMAGE:40052"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain 1NIB"
/note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not I; Site_2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 18.8%; Score 167; DB 5; Length 535;
Best Local Similarity 100.0%; Pred. No. 1e-78;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        721 ATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 ATGTGTGAGAACAAGAACTGAGGGAAAGCACCTTCCCCCACCCCAGACGTTTTTATCTGA 476

Qy        781 ATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 ATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAGTACT 416

Qy        841 CAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
              ||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 369
```

RESULT 3
AI813346/c
LOCUS AI813346 493 bp mRNA linear EST 21-DEC-1999
DEFINITION wj33g03.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2404660 3', mRNA sequence.
ACCESSION AI813346
VERSION AI813346.1 GI:5424561
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 493)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 879 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 447.
FEATURES Location/Qualifiers
source 1. .493
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2404660"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid12"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 12.5%; Score 111; DB 1; Length 493;
Best Local Similarity 100.0%; Pred. No. 2.6e-48;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        777 CTGAATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAG 836
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 CTGAATGCATACAAGGAGTCCTGAGGTGGTGATTTGGCCAGTGTTTTAACTTGTGACAAG 410

Qy        837 TACTCAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 887
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 TACTCAGGTGTGAGGACAAGAATGCAAATGGCTCTTCCTTGAGTGAAAGAA 359
```

RESULT 4
AA098865
LOCUS AA098865 130 bp mRNA linear EST 28-OCT-1996
DEFINITION zk84f02.s1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone IMAGE:489531 3', mRNA sequence.
ACCESSION AA098865
VERSION AA098865.1 GI:1645049
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 130)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 81.
FEATURES Location/Qualifiers
source 1. .130
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3804118"
/db_xref="taxon:9606"
/clone="IMAGE:489531"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."

ORIGIN

Query Match 10.3%; Score 91; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.7e-37;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        242 TCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCC 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCCGCCTACCTCGGCTACCCCGGGAACCGCTTCGAGCTGGTGGCGCTGATGGCGGATTCC 60

Qy        302 GTGCTCTCCGACAGCCCCGGCCCCACCTGGG 332
              |||||||||||||||||||||||||||||||
Db         61 GTGCTCTCCGACAGCCCCGGCCCCACCTGGG 91
```

RESULT 5
AA005293
LOCUS AA005293 206 bp mRNA linear EST 07-MAY-1997
DEFINITION zh93a11.r1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:428828 5', mRNA sequence.
ACCESSION AA005293
VERSION AA005293.1 GI:1448755
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 206)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 849 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 126.
FEATURES Location/Qualifiers
source 1. .206
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1328597"
/db_xref="taxon:9606"
/clone="IMAGE:428828"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen 1NFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 6.0%; Score 53; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 7.5e-17;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        383 CCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGA 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 CCGCTGGTGACCGCCCGGTGGAAGAAGTGGGGCTTCCAGCCGCGGCTAAAGGA 152


RESULT 6
R53538/c
LOCUS       R53538                  478 bp    mRNA    linear   EST 18-MAY-1995
DEFINITION  yg84c01.s1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:40052 3', mRNA sequence.
ACCESSION   R53538
VERSION     R53538.1  GI:815440
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 478)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1472
            High quality sequence stops: 402 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1472   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 402.
FEATURES             Location/Qualifiers
     source          1. .478
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:412593"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:40052"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares infant brain 1NIB"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                     I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'];
                     double-stranded cDNA was ligated to Hind III adaptors
                     (Pharmacia), digested with Not I and directionally cloned
                     into the Not I and Hind III sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             3.4%;  Score 30;  DB 7;  Length 478;
  Best Local Similarity   100.0%;  Pred. No. 0.00024;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        832 ACAAGTACTCAGGTGTGAGGACAAGAATGC 861
              ||||||||||||||||||||||||||||||
Db        431 ACAAGTACTCAGGTGTGAGGACAAGAATGC 402


RESULT 7
CNS02C0P/c
LOCUS       CNS02C0P                939 bp    DNA     linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            254D23 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL190546
VERSION     AL190546.1  GI:7828650
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
   PUBMED   10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
   PUBMED   10899143
REFERENCE   3  (bases 1 to 939)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .939
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="254D23"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG254CB12LP1~end : T7"
ORIGIN

  Query Match             3.0%;  Score 27;  DB 9;  Length 939;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        680 ACCCGCTTCTACCTGCCCAACTGTGAC 706
              |||||||||||||||||||||||||||
Db        503 ACCCGCTTCTACCTGCCCAACTGTGAC 477


RESULT 8
R85550
LOCUS       R85550                  316 bp    mRNA    linear   EST 14-AUG-1995
DEFINITION  yo38d07.s1 Soares adult brain N2b4HB55Y Homo sapiens cDNA clone
            IMAGE:180205 3', mRNA sequence.
ACCESSION   R85550
VERSION     R85550.1  GI:943956
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 316)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
```

```
            Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
            Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1578
            High quality sequence stops: 278
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1578   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 278.
FEATURES             Location/Qualifiers
     source          1. .316
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3826849"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:180205"
                     /sex="Male"
                     /dev_stage="55-year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares adult brain N2b4HB55Y"
                     /note="Organ: brain; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 53. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. The adult brain RNA was
                     provided by Dr. Donald H. Gilden. Tissue was acquired
                     17-18 hours after death which occurred in consequence of a
                     ruptured aortic aneurysm. RNA was prepared from a pool of
                     tissues representing the following areas of the brain:
                     frontal, parietal, temporal and occipital cortex from the
                     left and right hemispheres, subcortical white matter,
                     basal ganglia, thalamus, cerebellum, midbrain, pons and
                     medulla."
ORIGIN

  Query Match             2.5%;  Score 22;  DB 7;  Length 316;
  Best Local Similarity   100.0%;  Pred. No. 5.1;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        512 GCCTGGCTGCAGGCTCAGGGCG 533
              ||||||||||||||||||||||
Db        212 GCCTGGCTGCAGGCTCAGGGCG 233


RESULT 9
BZ869242
LOCUS       BZ869242                601 bp    DNA     linear   GSS 18-MAR-2003
DEFINITION  CH240_211K10.TJ CHORI-240 Bos taurus genomic clone CH240_211K10,
            genomic survey sequence.
ACCESSION   BZ869242
VERSION     BZ869242.1  GI:29096647
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE   1  (bases 1 to 601)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Chen,D., Riggs,F., de Jong,P.,
            Crawford,A.M. and McEwan,J.C.
  TITLE     Bovine BAC End Sequences from Library CHORI-240
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by AgResearch Ltd., New Zealand and The
            Institute of Genomic Research (TIGR), USA.
            Plate: 211  row: K  column: 10
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .601
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_211K10"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match             2.5%;  Score 22;  DB 8;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 5.3;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        744 GAAAGCACCTTCCCCCACCCCA 765
              ||||||||||||||||||||||
Db        229 GAAAGCACCTTCCCCCACCCCA 250


RESULT 10
AZ617122/c
LOCUS       AZ617122                441 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0448D17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0448D17 F, genomic survey sequence.
ACCESSION   AZ617122
VERSION     AZ617122.1  GI:11739312
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 441)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
```

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0448 row: D column: 17
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 441.

FEATURES Location/Qualifiers
source 1. .441
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0448D17"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 2.4%; Score 21; DB 8; Length 441;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 586 GAGAAAACAGCTGGTCCAGGC 606
||||||||||||||||||||
Db 276 GAGAAAACAGCTGGTCCAGGC 256

RESULT 11
BY590905/c
LOCUS BY590905 462 bp mRNA linear EST 15-DEC-2002
DEFINITION BY590905 RIKEN full-length enriched, adult inner ear Mus musculus cDNA clone F930014I02 3', mRNA sequence.
ACCESSION BY590905
VERSION BY590905.1 GI:26926087
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 462)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 22354683
PUBMED 12466851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Kirk W. Beisel (Boys Town National Research Hospital 555 North 30th Street Omaha,NE 68131 USA) whose assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .462
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930014I02"
/tissue_type="inner ear"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult inner ear"

ORIGIN

Query Match 2.4%; Score 21; DB 5; Length 462;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 748 GCACCTTCCCCCACCCCAGAC 768
||||||||||||||||||||

```
Db        262 GCACCTTCCCCCACCCCAGAC 242


RESULT 12
AQ294438
LOCUS       AQ294438               479 bp    DNA     linear   GSS 15-DEC-1998
DEFINITION  HS_3014_B2_H04_T7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3014 Col=8 Row=P, genomic survey
            sequence.
ACCESSION   AQ294438
VERSION     AQ294438.1  GI:4012021
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 479)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 3014  row: P  column: 8
            Class: BAC ends
            High quality sequence stop: 479.
FEATURES             Location/Qualifiers
     source          1. .479
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=3014 Col=8 Row=P"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
ORIGIN

  Query Match             2.4%;  Score 21;  DB 8;  Length 479;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        745 AAAGCACCTTCCCCCACCCCA 765
              |||||||||||||||||||||
Db        219 AAAGCACCTTCCCCCACCCCA 239


RESULT 13
AA823641
LOCUS       AA823641               494 bp    mRNA    linear   EST 17-FEB-1998
DEFINITION  vr69a10.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
            IMAGE:1125882 5', mRNA sequence.
ACCESSION   AA823641
VERSION     AA823641.1  GI:2893509
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:615218
            High quality sequence stop: 493.
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J x DBA/2J F1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1125882"
                     /tissue_type="embryo"
                     /dev_stage="2-cell"
                     /lab_host="DH10B"
                     /clone_lib="Knowles Solter mouse 2 cell"
                     /note="Organ: embryo; Vector: pBluescribe (modified);
                     Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
                     mRNA prepared from 13,500 2-cell stage embryos.  Primer:
                     SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'.  cDNAs
                     were cloned into the MluI/SalI sites of a modified
                     pBluescribe vector using commercial linkers (NEB).
                     Average insert size:  1.2 kb."
ORIGIN

  Query Match             2.4%;  Score 21;  DB 1;  Length 494;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        748 GCACCTTCCCCCACCCCAGAC 768
              |||||||||||||||||||||
Db        365 GCACCTTCCCCCACCCCAGAC 385


RESULT 14
BB628603
LOCUS       BB628603               657 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB628603 RIKEN full-length enriched, 16 days neonate cerebellum Mus
            musculus cDNA clone 9630018N03 5', mRNA sequence.
ACCESSION   BB628603
VERSION     BB628603.1  GI:16465993
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
```

Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

```
FEATURES             Location/Qualifiers
     source          1. .657
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="9630018N03"
                     /tissue_type="cerebellum"
                     /dev_stage="16 days neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 16 days neonate
                     cerebellum"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 20.0 and subtraction to Rot = 370.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I."
ORIGIN

  Query Match             2.4%;  Score 21;  DB 2;  Length 657;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        748 GCACCTTCCCCCACCCCAGAC 768
              ||||||||||||||||||||
Db         75 GCACCTTCCCCCACCCCAGAC 95


RESULT 15
CNS0413S/c
LOCUS       CNS0413S                803 bp    DNA     linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            073P11 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL269713
VERSION     AL269713.1  GI:7991606
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
   PUBMED   10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
   PUBMED   10899143
REFERENCE   3  (bases 1 to 803)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .803
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="073P11"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG073CH06SP1~end :
                     PUC-Ori"
ORIGIN

  Query Match             2.4%;  Score 21;  DB 9;  Length 803;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        710 CTAAATGACAGATGTGTGAGA 730
              |||||||||||||||||||||
Db        774 CTAAATGACAGATGTGTGAGA 754


Search completed: November  7, 2004, 06:03:22
Job time : 4554 secs
```